

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

**A Thesis Submitted for the Degree of PhD at the University of Warwick**

<http://go.warwick.ac.uk/wrap/67023>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.



# Well-quasi-ordering of Combinatorial Structures

by

**Aistis Atminas**

**Thesis**

Submitted to the University of Warwick

for the degree of

**Doctor of Philosophy**

**Mathematics Institute**

February 2015

THE UNIVERSITY OF  
**WARWICK**

# Contents

<b>List of Tables</b>	<b>iv</b>
<b>List of Figures</b>	<b>v</b>
<b>Acknowledgments</b>	<b>vii</b>
<b>Declarations</b>	<b>viii</b>
<b>Abstract</b>	<b>ix</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Basic Definitions and Conventions . . . . .	3
1.2 Partial orders and WQO . . . . .	3
1.3 Hereditary properties of partial orders . . . . .	5
1.4 Clique-width . . . . .	7
<b>I WQO for words, graphs and permutations</b>	<b>8</b>
<b>Chapter 2 Factorial languages</b>	<b>9</b>
2.1 Languages and automata . . . . .	9
2.2 Auxiliary results . . . . .	10
2.3 Main results . . . . .	13
2.4 An alternative approach . . . . .	18
2.5 Deciding WQO for factorial languages with binary alphabet . . . . .	19
2.6 Conclusion . . . . .	24
<b>Chapter 3 Graphs</b>	<b>28</b>
3.1 Induced Subgraph vs Labelled Induced Subgraph . . . . .	28
3.1.1 Higman’s lemma and Kruskal’s theorem . . . . .	29
3.1.2 Modular decomposition . . . . .	30

3.1.3	Minimal infinite antichains . . . . .	30
3.1.4	Modular decomposition . . . . .	32
3.1.5	Canonical decomposition of bipartite graphs . . . . .	33
3.2	An overview of graph classes . . . . .	34
3.2.1	$k$ -letter graphs . . . . .	34
3.2.2	$k$ -uniform graphs . . . . .	35
3.2.3	Monotone classes . . . . .	35
3.2.4	Other classes of graphs . . . . .	36
3.3	New results . . . . .	39
3.3.1	Subclasses of $P_6$ -free graphs . . . . .	39
3.3.2	$(P_k, Sun_2)$ -free bipartite graphs . . . . .	42
3.4	Conclusion . . . . .	44
<b>Chapter 4</b>	<b>Permutations and permutation graphs</b>	<b>45</b>
4.1	Definitions and structural tools . . . . .	45
4.2	Permutation Graphs Omitting $P_5$ and $K_\ell$ . . . . .	50
4.3	Permutation Graphs Omitting $P_6$ , $P_7$ or $P_8$ . . . . .	58
4.4	Conclusion . . . . .	62
<b>II</b>	<b>WQO and Clique-width</b>	<b>63</b>
<b>Chapter 5</b>	<b>WQO, Clique-width and The Speed of Hereditary Properties</b>	<b>64</b>
5.1	The speed of hereditary graph properties and the Bell number . . . . .	65
5.2	Preliminaries and preparatory results . . . . .	66
5.2.1	$(\ell, d)$ -graphs and sparsification . . . . .	66
5.2.2	Distinguishing number $k_\chi$ . . . . .	69
5.3	Structure of minimal classes above Bell . . . . .	70
5.3.1	Infinite distinguishing number . . . . .	70
5.3.2	Finite distinguishing number . . . . .	74
5.4	Decidability of the Bell number . . . . .	83
5.5	Deciding WQO for hereditary classes with finite distinguishing number . . . . .	88
5.5.1	Classes below the Bell number . . . . .	88
5.5.2	Classes above the Bell number with finite distinguishing number . . . . .	90
5.6	Conclusion . . . . .	92

<b>Chapter 6</b>	<b>Subquadratic properties</b>	<b>93</b>
6.1	Notations and preparatory results . . . . .	94
6.1.1	Infinite antichains . . . . .	94
6.1.2	Excluded graphs in subquadratic classes . . . . .	95
6.1.3	Long paths in graphs in subquadratic classes . . . . .	96
6.1.4	Some bounds on the number $Z(s, t, q)$ . . . . .	97
6.2	Main result: WQO classes of bounded treewidth . . . . .	99
6.2.1	Plan of the proof . . . . .	100
6.2.2	Rake subgraphs in graphs of large tree-width . . . . .	101
6.2.3	From rake subgraphs to dense rake subgraphs . . . . .	101
6.2.4	Dense rake subgraphs . . . . .	104
6.2.5	Canonical graphs and bicliques in graphs of large tree-width .	105
6.2.6	Proof of Theorem 6.2.1 . . . . .	106
6.3	Deciding WQO for subquadratic classes . . . . .	106
6.4	Byproduct: Computing small bicliques in graphs without long in- duced paths . . . . .	107
6.5	Conclusion . . . . .	109
<b>Chapter 7</b>	<b>Classes of unbounded clique-width</b>	<b>111</b>
7.1	Universal graphs and universal permutations . . . . .	112
7.1.1	Universal 321-avoiding permutations . . . . .	113
7.1.2	Universal split permutation and universal bichain graphs . .	120
7.1.3	Z-grid . . . . .	124
7.2	Minimality of classes of bichain and split permutation graphs . . . .	128
7.2.1	Bipartite permutation graphs and the $X$ -grid . . . . .	128
7.2.2	Pivoting bichain graphs into bipartite permutation graphs . .	132
7.2.3	Proof of minimality . . . . .	133
7.3	Canonical antichain . . . . .	134
7.4	Conclusion . . . . .	139

# List of Tables

2.1	Minimal periodic infinite antichains of period 1 and 2 . . . . .	20
2.2	Minimal periodic infinite antichains of period 3 . . . . .	20
2.3	Tripples $(\alpha_1, \alpha_2, \alpha_3)$ destroying antichains of period at most 3 . . . .	24

# List of Figures

1.1	Graphs $H_k$ (left) and $Sun_4$ (right) . . . . .	5
2.1	A colored path on $P_5$ . . . . .	26
3.1	The graph $Sun_2$ . . . . .	42
4.1	WQO results for classes of graphs and permutation graphs avoiding paths and cliques . . . . .	46
4.2	The impossible configuration for a simple permutation in Proposition 4.2.1. . . . .	51
4.3	The division of $\pi$ used in Proposition 4.2.2 (on the left) and the final contradictions in the proof (in the centre and on the right) . . . . .	53
4.4	Propagating a side of a hull in $\mathfrak{H}$ , and its corresponding propagation tree. . . . .	55
4.5	Two independent bad rectangles . . . . .	57
4.6	The first three elements and the general structure of the infinite antichain $A$ in Grid $(\oplus^{21} \oplus^{21})$ . . . . .	60
4.7	The first three elements of an infinite antichain in Grid $(\oplus^{21} \overset{Av(21)}{\oplus^{21}})$ . . . . .	61
5.1	The universal graphs . . . . .	71
5.2	Some small graphs . . . . .	72
6.1	Graphs $H'_i$ and $H''_i$ . . . . .	94
6.2	1-dense 9-rake . . . . .	100
7.1	The universal bipartite permutation graph $X_{n,n}$ with a labeling. . . . .	114
7.2	The permutation $\rho_n$ corresponding to the graph in Figure 7.1. . . . .	115
7.3	A direct reading off the permutation $\rho_n$ from the graph. . . . .	116
7.4	Applying inductive hypothesis for Theorem 7.1.1 . . . . .	117
7.5	Examples illustrating our inductive proof of Theorem 7.1.1 . . . . .	118

7.6	Universal threshold graph for $n = 5$ (the oval represents a clique) . . .	121
7.7	Graph $Z_{6,7}$ without the "diagonal" edges . . . . .	126
7.8	Graphs $G_4^1 = \phi(\rho_4^1)$ (left) and $G_4^2 = \phi(\rho_4^2)$ (right) without "diagonal" and "clique" edges . . . . .	126
7.9	$Z_{11,72}^*$ contains the universal bichain graph $\phi((\rho_4^1 \rho_4^2)^4)$ . . . . .	127
7.10	Embedding a graph into a $Z$ -grid with fewer columns . . . . .	128
7.11	Graph $X_{6,6}$ . . . . .	129
7.12	Graphs $Y_{7,5}$ (left) and $Y_{7,5}^+$ (right) . . . . .	130
7.13	The graph $X_{6,6}$ contains the graph $Y_{3,3}^+$ (left) and the graph $Y_{6,5}^+$ contains the graph $X_{3,3}$ (right) . . . . .	130
7.14	The graph $Z_{7,5}^+$ with "diagonal" edges omitted (left) transforms into the graph $Y_{7,5}^+$ (right) by pivoting on the edges $cd, ef, gh$ . . . . .	133
7.15	The graph $S_7$ . . . . .	134
7.16	Reading $Z_{n,k}$ as a $k$ -letter graph . . . . .	136



# Acknowledgments

First and foremost, I would like to thank my PhD supervisor Dr. Vadim V. Lozin for his extraordinary patience and for all the help, advice and encouragement he has given me over the past three years. I would also like to thank him for his guidance and expertise in the research on which we collaborated. These thanks also extend to all my co-authors.

I would like to thank Dr. Ross Lawther, Dr. Dennis Barden and Dr. Ruth Williams who were nurturing my interest and taking exceptionally good care of me during my undergraduate studies at the University of Cambridge. My thanks extend to many great lecturers, supervisors and friends at Cambridge whose enthusiasm sustained and further developed my interest in the subject.

I would also like to thank Dr. Romualdas Kašuba and my high school math teacher Mrs. Irena Bagdonienė who inspired me to take part in mathematical competitions and guided me through the early stages of my career in mathematics.

This thesis would not be possible without support from my family. I am grateful for my parents and my sister who have given me selfless, unconditional love and support. My thanks also go to Chenxi Li for her love and support.

Finally, I would like to thank all the staff and students of the Centre for Discrete Mathematics and its Applications (DIMAP) and Mathematics Institute at the University of Warwick. This research was generously funded by DIMAP, which was partially supported by EPSRC Science and Innovation Award EP/D063191/1.

# Declarations

All the work in this thesis is joint work with my thesis supervisor Vadim Lozin. This work has not been submitted for a degree at another university.

- Chapter 2 is based on [Atminas et al., 2013b] which was joint work with V. Lozin and M. Moshkov.
- Chapter 3 is based on [Atminas and Lozin, 2014] which was joint work with V. Lozin.
- Chapter 4 is based on [Atminas et al., 2013a] which was joint work with V. Lozin, R. Brignal, N. Korpelainen and V. Vatter.
- Chapter 5 is based on [Atminas et al., 2014a] which was joint work with V. Lozin, I. Foniok and A. Collins.
- Chapter 6 is based on [Atminas et al., 2014b] which was joint work with V. Lozin and I. Razgon.
- Chapter 7 is based on three independent works:
  - Sections 7.1.1 and 7.1.2 are based on [Atminas et al., 2013c] which was joint work with S. Kitaev and A. Valyuzhenic,
  - Sections 7.1.3 and 7.2 are based on my recent unpublished joint work with R. Brignal, N. Korpelainen, V. Lozin and J. Stacho
  - Section 7.3 is based on my recent unpublished joint work with V. Lozin.

# Abstract

In this work we study the notion of well-quasi-ordering for various partial orders and its relation to some other notions, such as clique-width. In particular, we prove decidability of well-quasi-ordering for factorial languages, subquadratic properties of graphs and classes of graphs with finite distinguishing number. In addition, we reveal some new classes of graphs and permutations which are or are not well-quasi-ordered. We also prove that subquadratic properties or classes of graphs with finite distinguishing number that are well-quasi-ordered have bounded clique-width and we identify two new minimal classes of graphs of unbounded clique-width.

# Chapter 1

## Introduction

Well-quasi-ordering (wqo) is a highly desirable property and frequently discovered concept in mathematics and theoretical computer science [Finkel and Schnoebelen, 2001, Kruskal, 1972]. One of the most remarkable recent results in this area is the proof of Wagner’s conjecture stating that the set of all finite graphs is well-quasi-ordered by the minor relation [Robertson and Seymour, 2004]. However, the subgraph or induced subgraph relation is not a well-quasi-order. Other examples of important relations that are not well-quasi-orders are pattern containment relation on permutations [Spielman and Bóna, 2000], embeddability relation on tournaments Cherlin and Latka [2000], minor ordering of matroids [Hine and Oxley, 2010], factor (contiguous subword) relation on words [de Luca and Varricchio, 1994]. On the other hand, each of these relations may become a well-quasi-order under some additional restrictions. In this work, we study restrictions given in the form of obstructions, i.e. minimal excluded (“forbidden”) elements. The fundamental problem of our interest is the following:

*Problem 1.* Given a partial order  $P$  and a *finite* set of obstructions  $Z$ , determine if the set of elements of  $P$  containing no elements from  $Z$  forms a well-quasi-order.

This problem was studied for the induced subgraph relation on graphs [Korpelainen and Lozin, 2011a], the pattern containment relation on permutations [Brignall et al., 2008], the embeddability relation on tournaments [Cherlin and Latka, 2000], the minor ordering of matroids [Hine and Oxley, 2010]. However, the decidability of this problem has been shown only for one or two forbidden elements (graphs, permutations, tournaments, matroids).

Part I is devoted to the study of Problem 1 for words, permutations and graphs. In Chapter 2 we show that the problem for factorial languages can be solved in polynomial time. This positive result for words motivates to look at peri-

odic constructions of antichains for other partial orders, in particular in Chapter 2.6 we observe some promising ways of extending this result to the graphs with labelled induced subgraph relation. We then conjecture that answering Problem 1 for labelled induced subgraph relation is equivalent to induced subgraph relation. In other words, we conjecture that a finitely defined class of graphs is well-quasi-ordered by induced subgraph relation if and only if it is well-quasi-ordered by labelled induced subgraph relation. In Chapter 3, we verify this conjecture for all classes known to be well-quasi-ordered by induced subgraph relation. We also contribute to the list of partial results for classes defined by two forbidden elements for graphs in the end of Chapter 3 and permutations in Chapter 4.

In part II, we investigate the relation between the notion of well-quasi-ordering and the notion of clique-width. This relation has been observed by Daligault, Rao, and Thomassé [2010]:

**Conjecture 1.** Every hereditary class which is well-quasi-ordered by the induced subgraph relation is of bounded clique-width.

In Chapter 5 and Chapter 6 we verify this conjecture for classes of finite distinguishing number and subquadratic properties. As a byproduct, we also show that the task to test well-quasi-ordering by the induced subgraph relation for these two families of graph properties is decidable. Chapter 5 also uncovers an interesting relation between the notion of well-quasi-ordering and the speed of hereditary properties. In fact, we show that all classes with speed below Bell number are well-quasi-ordered and finitely defined above Bell number with finite distinguishing number are not well-quasi-ordered. Deciding well-quasi-ordering for the classes with infinite distinguishing number remains an open research question.

Conjecture 1 could also be rewritten into equivalent statement saying that every minimal hereditary class of unbounded clique-width contains an infinite antichain. We conjecture that even more is true, that such classes contain a canonical antichain. In Section 7.2 we prove that the class of split permutation graphs and its analog a class of bichain graphs are minimal classes of unbounded clique-width and in Section 7.3 we prove that these classes contain a canonical antichain with respect to labelled induced subgraph relation. Section 7.1 serves as a preparatory result constructing a universal graphs for these classes.

The rest of this chapter contains basic definitions and notations used in this work and some preliminary information concerning hereditary classes and the notion of well-quasi-ordering.

## 1.1 Basic Definitions and Conventions

All graphs in this thesis are *simple*, i.e. undirected, without loops and multiple edges. However, we allow loops in Chapter 5 in some auxiliary graphs, called "density graphs" and denoted usually by  $H$ , that are used to represent the global structure of our graph classes. The vertex set and the edge set of a graph  $G$  are denoted by  $V(G)$  and  $E(G)$  respectively. Given two disjoint subsets  $U \subset V(G)$  and  $W \subset V(G)$ , we say that  $U$  is *complete* to  $W$  if every vertex of  $U$  is adjacent to every vertex of  $W$ . Also, we say that  $U$  is *anticomplete* to  $W$  if no vertex of  $U$  is adjacent to a vertex of  $W$ .

For a subset  $U \subseteq V(G)$ , we denote by  $G[U]$  the subgraph of  $G$  *induced* by  $U$ , i.e. the graph with vertex set  $U$  in which two vertices are adjacent if and only if they are adjacent in  $G$ . We say that a graph  $G$  contains a graph  $H$  as an *induced subgraph* if  $H$  is isomorphic to an induced subgraph of  $G$ . If  $G$  contains no induced subgraphs isomorphic to  $H$ , we say that  $G$  is  *$H$ -free*.

For a bipartite graph  $G = (A, B, E)$  given together with a bipartition of its vertex set onto two independent sets  $A$  and  $B$ , we denote by  $\tilde{G}$  the *bipartite complement* of  $G$ , i.e. the bipartite graph obtained from  $G$  by complementing the edges between  $A$  and  $B$ .

We denote by  $K_n$ ,  $P_n$  and  $C_n$  the complete graph, the chordless path and the chordless cycle on  $n$  vertices, respectively. Also,  $S_{1,2,3}$  is the tree with exactly three vertices of degree 1 being of distance 1, 2, 3 from the only vertex of degree 3, and a *diamond* is the graph obtained from  $K_4$  by deleting an edge. By  $Sun_4$  we denote the graph represented in Figure 1.1(right), and by  $Sun_i$  ( $1 \leq i \leq 4$ ), the graph obtained from  $Sun_4$  by deleting  $4 - i$  vertices of degree 1. In particular,  $Sun_2$  is the graph represented in Figure 3.1 in Section 3.3.2. Given two graphs  $G$  and  $H$ , we denote by  $G + H$  the disjoint union of  $G$  and  $H$ .

## 1.2 Partial orders and WQO

For a set  $A$  we denote by  $A^2$  the set of all ordered pairs of (not necessarily distinct) elements from  $A$ . A *binary relation* on  $A$  is a subset of  $A^2$ . If a binary relation  $\mathcal{R} \subset A^2$  is

- *reflexive*, i.e.  $(a, a) \in \mathcal{R}$  for each  $a \in A$ ,
- *transitive*, i.e.  $(a, b) \in \mathcal{R}$  and  $(b, c) \in \mathcal{R}$  imply  $(a, c) \in \mathcal{R}$ ,

then  $\mathcal{R}$  is a *quasi-order* (also known as *pre-order*). If additionally  $\mathcal{R}$  is

- *antisymmetric*, i.e.  $(a, b) \in \mathcal{R}$  and  $(b, a) \in \mathcal{R}$  imply  $a = b$ ,

then  $\mathcal{R}$  is a *partial order*.

We say that two elements  $a, b \in A$  are comparable with respect to  $\mathcal{R}$  if either  $(a, b) \in \mathcal{R}$  or  $(b, a) \in \mathcal{R}$ . A set of pairwise comparable elements of  $A$  is called a *chain* and a set of pairwise incomparable elements of  $A$  is called an *antichain*.

A quasi-ordered set is *well-quasi-ordered* if it contains

- neither infinite strictly decreasing chains, in which case we say that the set is *well-founded*,
- nor infinite antichains.

All examples of quasi-orders in this thesis will be antisymmetric (i.e. partial orders) and well-founded, in which case well-quasi-orderability is equivalent to the non-existence of infinite antichains.

*Examples.*

- (1) Let  $A$  be the set of all finite simple (i.e. undirected, without loops and multiple edges) graphs. If a graph  $H \in A$  can be obtained from a graph  $G \in A$  by a (possibly empty) sequence of
  - vertex deletions, then  $H$  is an *induced subgraph* of  $G$ ,
  - vertex deletions and edge deletions, then  $H$  is a *subgraph* of  $G$ ,
  - vertex deletions, edge deletions and edge contractions, then  $H$  is a *minor* of  $G$ ,
  - vertex deletions and edge contractions, then  $H$  is an *induced minor* of  $G$ .

According to the celebrated result of Robertson and Seymour [2004] the minor relation on the set of graphs is a well-quasi-order. However, this is not the case for the subgraph, induced subgraph and induced minor relation. Indeed, it is not difficult to see that the set of all chordless cycles  $C_3, C_4, C_5, \dots$  creates an infinite antichain with respect to both subgraph and induced subgraph relations, and the complements of the cycles form an infinite antichain with respect to the induced minor relation. Besides, the set of so-called *H*-graphs (i.e. graphs represented in Figure 1.1(left)) also forms an infinite antichain with respect to subgraph and induced subgraph relations.

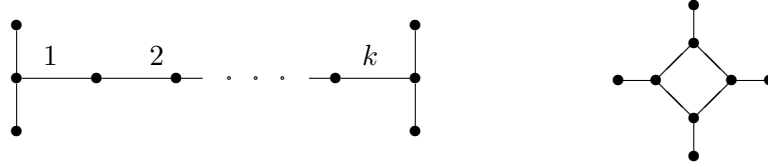


Figure 1.1: Graphs  $H_k$  (left) and  $Sun_4$  (right)

- (2) Let  $A$  be the set of all finite permutations. We say that a permutation  $\pi \in A$  of  $n$  elements is contained in a permutation  $\rho \in A$  of  $k$  elements ( $n \leq k$ ) as a *pattern*, if  $\pi$  can be obtained from  $\rho$  by deleting some (possibly none) elements and renaming the remaining elements consecutively in the increasing order. Obviously, the pattern containment relation is a well-founded partial order. However, whether it is a quasi-order is not an obvious fact. Finding an infinite antichain of permutations becomes much easier if we associate to each permutation its permutation graph. Let  $\pi$  be a permutation on the set  $N = \{1, 2, \dots, n\}$ . The permutation graph  $G_\pi$  of  $\pi$  is the graph with vertex set  $N$  in which two vertices  $i$  and  $j$  are adjacent if and only if they form an inversion in  $\pi$  (i.e.  $i < j$  and  $\pi(i) > \pi(j)$ ). It is not difficult to see that if  $\rho$  contains  $\pi$  as a pattern, then  $G_\rho$  contains  $G_\pi$  as an induced subgraph. Therefore, if  $G_1, G_2, \dots$  is an infinite antichain of permutation graphs with respect to the induced subgraph relation, then the corresponding permutations form an infinite antichain with respect to the pattern containment relation. Since the  $H$ -graphs (Figure 1.1) are permutation graphs (which is easy to see), we conclude that the pattern containment relation on permutations is not a well-quasi-order.

- (3) Let  $A$  be the set of all finite words in a finite alphabet. A word  $\alpha \in A$  is said to be a *factor* of a word  $\beta \in A$  if  $\alpha$  can be obtained from  $\beta$  by omitting a (possibly empty) suffix and prefix. If the alphabet contains at least two symbols, say 1 and 0, the factor containment relation is not a well-quasi-order, since it necessarily contains an infinite antichain, for instance, 010, 0110, 01110, etc.

### 1.3 Hereditary properties of partial orders

Let  $(A, \mathcal{R})$  be a well-founded partial order. A property on  $A$  is a subset of  $A$ . A property  $P \subseteq A$  is *hereditary* (with respect to  $\mathcal{R}$ ) if  $x \in P$  implies  $y \in P$  for every



$y \in A$  such that  $(y, x) \in \mathcal{R}$ . Hereditary properties are also known as lower ideals or downward closed sets.

*Examples.*

- If  $A$  is the set of all finite graphs and  $\mathcal{R}$  is the minor relation, then a hereditary property on  $A$  is known as a *minor-closed* class of graphs.
- If  $A$  is the set of all finite graphs and  $\mathcal{R}$  is the subgraph relation, then a hereditary property on  $A$  is known as a *monotone* class of graphs.
- If  $A$  is the set of all permutations and  $\mathcal{R}$  is the pattern containment relation, then a hereditary property on  $A$  is known as a *pattern* class or *pattern avoiding* class.
- If  $A$  is the set of all words in a finite alphabet and  $\mathcal{R}$  is the factor containment relation, then a hereditary property on  $A$  is known as a *factorial language*.

The word “avoiding” used in the terminology of permutations suggests that a hereditary property can be described in terms of “forbidden” elements. To better explain this idea, let us introduce the following notation: given a set  $Z \subseteq A$ , we denote

$$Free(Z) := \{a \in A \mid (z, a) \notin \mathcal{R} \forall z \in Z\}.$$

Obviously, for any  $Z \subseteq A$ , the set  $Free(Z)$  is hereditary. On the other hand, for any hereditary property  $P \subseteq A$  there is a unique minimal set  $Z \subseteq A$  such that  $P = Free(Z)$ . We call  $Z$  the set of *forbidden elements* for  $Free(Z)$  and observe that a minimal set of forbidden elements is necessarily an antichain.

*Examples.*

- Since the minor relation on graphs contains no infinite antichains, any minor-closed class of graphs can be described by a finite set of forbidden minors. In particular, for the class of planar graphs the set of minimal forbidden minors consists of  $K_5$ , the complete graph on 5 vertices, and  $K_{3,3}$ , the complete bipartite graph with 3 vertices in each part.
- The set of minimal forbidden permutations for a pattern avoiding class is also known as the *base* of the class.
- The set of minimal forbidden words for a factorial language is also known as the *antidictionary* of the language.

In the above notation, the problem of our interest can be stated as follows:

*Problem 1.3.1.* Given a finite set  $Z \subset A$ , determine if  $Free(Z)$  is well-quasi-ordered with respect to  $\mathcal{R}$ .

This question is not applicable to the minor relation on graphs, since this relation is a well-quasi-order. For hereditary properties of graphs with respect to the subgraph relation, Problem 1.3.1 has a simple solution which is due to Ding [1992]: a monotone class of graphs is well-quasi-ordered by the subgraph relation if and only if it contains finitely many cycles and finitely many  $H$ -graphs. Therefore, if  $Z$  is finite, then  $Free(Z)$  is well-quasi-ordered with respect to the subgraph relation if and only if  $Z$  includes a chordless path (or an induced subgraph of a chordless path), because otherwise  $Free(Z)$  contains infinitely many cycles.

For other relations, such as the induced subgraph relation on graphs or pattern containment relation on permutations, only partial results are available, where  $Z$  contains one or two elements (see e.g. [Atkinson et al., 2002, Korpelainen and Lozin, 2011a]). Whether this problem is decidable for larger numbers of forbidden elements is an open question which we approach in the part I. In section 2 we study the Problem 1.3.1 for factorial languages and show that the problem is efficiently solvable for any finite set  $Z$ .

## 1.4 Clique-width

The notion of clique-width of a graph was introduced in [Courcelle et al., 1993] and is defined as the minimum number of labels needed to construct the graph using the following four graph operations:

1. Creation of a new vertex  $v$  with label  $i$  (denoted  $i(v)$ )
2. Disjoint union of two labelled graphs  $G$  and  $H$  (denoted  $G \oplus H$ )
3. Joining by an edge every vertex labelled  $i$  to every vertex labelled  $j$  (denoted  $\eta_{i,j}$ ), for some  $i \neq j$
4. Renaming label  $i$  to label  $j$  (denoted  $\rho_{i \rightarrow j}$ )

The importance of clique-width is due to the fact that many computational problems that are  $NP$ -hard in general can be solved in polynomial time in the classes of graphs of bounded clique-width [Arnborg and Proskurowski, 1989, Courcelle et al., 2000]. In the second part of the dissertation we will study the relation between well-quasi-ordering and clique-width and we will obtain some evidence that well-quasi-ordered classes have bounded clique-width.

## Part I

# WQO for words, graphs and permutations

## Chapter 2

# Factorial languages

A language is *factorial* if it is closed under taking factors (i.e. contiguous subwords). Every factorial language can be described by an antidictionary, i.e. a minimal set of forbidden factors. In this chapter we show that our fundamental problem can be solved efficiently for factorial languages, i.e. we show that one can decide in polynomial time whether a factorial language given by a finite antidictionary is well-quasi-ordered under the factor containment relation.

In Section 2.1, we review the result from [Crochemore et al., 1998] which allows us to represent a factorial language defined by finite antidictionary in the form of a finite deterministic automaton. In Sections 2.2 and 2.3 we use this representation to solve our problem. In Sections 2.4 and 2.5 we explore the periodicity of the antichains we obtained. We conclude this chapter in Section 2.6 by discussing possible extensions of our result to other combinatorial structures.

### 2.1 Languages and automata

Let  $k \geq 2$  be a natural number and  $E_k = \{0, 1, \dots, k-1\}$  be an alphabet. A *finite deterministic automaton* over  $E_k$  is a triple  $A = (G, q_0, Q)$ , where

- $G$  is a finite directed graph, possibly with multiple edges and loops, in which the edges are labeled with letters from  $E_k$  in such a way that any two edges leaving the same node have different labels,
- $q_0$  is a node of  $G$ , called the *start node*, and
- $Q$  is a nonempty set of nodes of  $G$ , called the *terminal nodes*.

A directed path in  $G$  is any sequence  $v_1, e_1, \dots, v_m, e_m, v_{m+1}$  of nodes  $v_i$  and edges  $e_j$  such that for each  $j = 1, \dots, m$ , the edge  $e_j$  is directed from  $v_j$  to  $v_{j+1}$ . We

emphasize that both nodes and edges can appear in such a path repeatedly.

With each directed path  $\tau$  in the graph  $G$  we associate a word over  $E_k$  by reading the labels of the edges of  $\tau$  (listed along the path) and denote this word by  $w(\tau)$ . A directed path in  $G$  will be called an  $A$ -path if it starts at the node  $q_0$  and ends at a terminal node.

Let  $\alpha$  be a word over  $E_k$ . We say that an automaton  $A = (G, q_0, Q)$  accepts  $\alpha$  if there is an  $A$ -path  $\tau$  such that  $w(\tau) = \alpha$ . The set of all words accepted by  $A$  is called the language accepted (or recognized) by  $A$  and this language is denoted  $L(A)$ . It is well-known that the set of languages accepted by finite deterministic automata are precisely the regular languages.

The following result was proved in [Crochemore et al., 1998].

**Theorem 2.1.1.** *Given a set  $Z$  of  $n$  words over  $E_k$ , in time  $O(nk)$  one can construct a finite deterministic automaton  $A$  such that  $L(A)$  coincides with the factorial language  $Free(Z)$ .*

We call an automaton  $A = (G, q_0, Q)$  *reduced* if for each node of  $G$  there exists an  $A$ -path containing this node. It is not difficult to see that any finite automaton can be transformed into an equivalent (i.e. accepting the same language) reduced automaton in polynomial time. This observation together with Theorem 2.1.1 reduce Problem 1.3.1 to the following one:

*Problem 2.1.1.* Given a finite deterministic automaton  $A$ , determine if  $L(A)$  is well-quasi-ordered with respect to the factor containment relation.

In the next two sections, we give an efficient solution to this problem.

## 2.2 Auxiliary results

Given a word  $\alpha$ , we denote by  $|\alpha|$  the length of  $\alpha$ , i.e. the number of letters in  $\alpha$ . Also,  $\alpha^i$  denotes concatenation of  $i$  copies of  $\alpha$  and is called the  $i$ -th power of  $\alpha$ .

A word  $\alpha = \alpha_1 \dots \alpha_n$  is called a *periodic word with period  $p$*  if

- either  $p \geq n$
- or  $p < n$  and  $\alpha_i = \alpha_{i+p}$  for  $i = 1, \dots, n - p$ .

A word  $\gamma$  will be called a *left extension of a power of a word  $\alpha$*  if  $\gamma$  can be represented in the form  $\sigma\alpha^i$ , where  $\sigma$  is a suffix of  $\alpha$ . Similarly,  $\gamma$  will be called a *right extension of a power of  $\alpha$*  if  $\gamma$  can be represented in the form  $\alpha^i\sigma$ , where  $\sigma$  is a prefix of  $\alpha$ . Directly from the definition we obtain the following conclusion.

**Lemma 2.2.1.** *A word  $\gamma$  is a left extension of a power of  $\alpha$  if and only if the word  $\gamma\alpha$  is a periodic word with period  $|\alpha|$ . A word  $\gamma$  is a right extension of a power of  $\alpha$  if and only if the word  $\alpha\gamma$  is a periodic word with period  $|\alpha|$ .*

Now we prove a number of further auxiliary results.

**Lemma 2.2.2.** *Let  $\gamma, \alpha, \delta$  be words such that either  $\gamma$  is a left extension of a power of  $\alpha$  or  $\delta$  is a right extension of a power of  $\alpha$ . Then the set  $\{\gamma\alpha^i\delta : i = 0, 1, 2, \dots\}$  is a chain, i.e. any two words in this set are comparable.*

*Proof.* Let  $\gamma\alpha^i\delta$  and  $\gamma\alpha^j\delta$  be two words with  $i < j$ . If  $\gamma$  is a left extension of a power of  $\alpha$ , then the word  $\gamma$  can be represented as  $\sigma\alpha^k$  where  $\sigma$  is a suffix of  $\alpha$ . Therefore, the word  $\gamma\alpha^i\delta$  can be obtained from the word  $\gamma\alpha^j\delta$  by removing a prefix of length  $(j-i)|\alpha|$ . Similarly, if  $\delta$  is a right extension of a power of  $\alpha$ , then the word  $\gamma\alpha^i\delta$  can be obtained from the word  $\gamma\alpha^j\delta$  by removing a suffix of length  $(j-i)|\alpha|$ .  $\square$

**Lemma 2.2.3.** *Let  $\gamma, \alpha, \delta$  be words such that  $\gamma$  is not a left extension of a power of  $\alpha$ , and  $\delta$  is not a right extension of a power of  $\alpha$ . Then the set of words  $\{\gamma\alpha^i\delta : i = 1, 2, \dots\}$  is an infinite antichain.*

*Proof.* Let  $\alpha = a_1a_2\dots a_p$  be a word of length  $p$ . Suppose there is a contiguous embedding  $\phi : \gamma\alpha^j\delta \rightarrow \gamma\alpha^k\delta$ . Then the factor  $\alpha^j$  of the first word is embedded into the factor  $\alpha^k$  of the second word. Denote the factor of the second word  $\alpha^k = a_1a_2\dots a_{kp}$  where  $a_i = a_{i(\bmod p)}$ . Suppose the first letter  $a_1$  is mapped by  $\phi$  to  $a_{h+1}$ . Then,  $a_m = a_{m+h}$  for  $m = 1, 2, \dots, p$ . Now  $a_m = a_{m+h(\bmod p)} = a_{m+2h(\bmod p)} \dots$ . We define  $p' = \gcd(h, p)$ , and claim that  $\alpha$  is a periodic word of period  $p'$ . Indeed, by Bézout's identity we have  $p' = xh - yp$  for some positive integers  $x$  and  $y$ . Hence  $a_m = a_{m+xh(\bmod p)} = a_{m+p'}$  for all  $m = 1, 2, \dots, (p - p')$ . This shows that  $\alpha$  is periodic word with period  $p'$ .

Define  $\alpha' = a_1a_2\dots a_{p'}$ , and write  $\alpha = (\alpha')^{p \setminus p'}$ . Notice that  $h = zp'$ , for some nonnegative integer  $z$ . If  $z \neq 0$ , we get that the factor of the first word  $\gamma$  is embedded to the suffix of  $\gamma(\alpha')^z$  from the second word. Adding extra powers of  $(\alpha')^z$  to both words we get that  $\gamma(\alpha')^{2z}$  is a suffix of  $\gamma(\alpha')^z$ ,  $\gamma(\alpha')^{3z}$  is a suffix of  $\gamma(\alpha')^{2z}$ ,  $\dots$ ,  $\gamma(\alpha')^{sz}$  is a suffix of  $\gamma(\alpha')^{(s-1)z}$ . So we get that  $\gamma$  is a suffix of  $\gamma(\alpha')^z$  which is a suffix of  $\gamma(\alpha')^{2z}$  which is a suffix of  $\gamma(\alpha')^{3z}$ ,  $\dots$ , which is a suffix of  $\gamma(\alpha')^{sz}$ . So  $\gamma$  is a suffix of  $\gamma(\alpha')^{sz}$ . For  $s$  larger than  $\frac{|\gamma|}{|(\alpha')^z|}$  we get that  $\gamma$  is a suffix of  $(\alpha')^{sz}$  which means that it is a left extension of a power of  $\alpha'$  and hence a left extension of a power of  $\alpha$ . A contradiction concludes that  $z = 0$ , hence  $h = 0$  and that the first word is the prefix of the second.

So consider the case when the first word  $\gamma\alpha^j\delta$  is the prefix of the second  $\gamma\alpha^k\delta$ . Then, in particular  $\delta$  from the first word is mapped to the prefix of  $(\alpha)^{(k-j)}\delta$  from the second word. Now, adding the powers of  $\alpha^{(k-j)}$ , we get that  $\alpha^{(k-j)}\delta$  is a prefix of  $\alpha^{2(k-j)}\delta$ ,  $\alpha^{2(k-j)}\delta$  is a prefix of  $\alpha^{3(k-j)}\delta$ , ...,  $\alpha^{(s-1)(k-j)}\delta$  is a prefix of  $\alpha^{s(k-j)}\delta$ . So  $\delta$  is a prefix of  $(\alpha)^{(k-j)s}\delta$ . If  $k - j > 0$  then we choose  $s$  to be larger than  $\frac{|\delta|}{|(\alpha)^{(k-j)}|}$  and we get that  $\delta$  is a prefix of  $\alpha^{(j-k)s}$ . But this means that  $\delta$  is a right extension of a power of  $\alpha$ . A contradiction. So  $k = j$  and hence no two different words are comparable in the set  $\{\gamma\alpha^i\delta : i = 1, 2, \dots\}$ , concluding that this set is an antichain.  $\square$

**Lemma 2.2.4.** *Let  $\gamma, \alpha, \delta, \beta$  be words. If  $\delta$  is not a right extension of a power of  $\alpha$ , then the word  $\gamma\alpha^{|\beta|}\delta$  is not a left extension of a power of  $\beta$ .*

*Proof.* Let us assume the contrary, i.e. assume that  $\gamma\alpha^{|\beta|}\delta$  is a left extension of a power of  $\beta$ . Then, by Lemma 2.2.1, the word  $\varepsilon = \gamma\alpha^{|\beta|}\delta\beta$  is a periodic word with period  $|\beta|$  and, therefore, with period  $|\beta||\alpha|$ . Since  $\alpha^{|\beta|}$  is a factor of  $\varepsilon$  and  $|\beta||\alpha|$  is a period of  $\varepsilon$ , the word  $\varepsilon$  is a factor of the word  $\alpha^{|\beta|p}$  for a large enough natural number  $p$ . Hence,  $\varepsilon$  is a periodic word with period  $|\alpha|$ . As a result,  $\alpha\delta$  also is a periodic word with period  $|\alpha|$ . However, this is impossible by Lemma 2.2.1, since  $\delta$  is not a right extension of a power of  $\alpha$ .  $\square$

**Lemma 2.2.5.** *Let  $\alpha_1, \alpha_2, \gamma, \sigma, \delta$  be words. If  $\sigma$  is a suffix of  $\alpha_1$  and  $\delta$  is a prefix of  $\alpha_2$ , then the set  $S = \{\sigma\alpha_1^i\gamma\alpha_2^j\delta \mid i, j \in \mathbb{N}\}$  is well-quasi-ordered by the factor containment relation.*

*Proof.* Let  $\varepsilon_1 = \sigma\alpha_1^{i_1}\gamma\alpha_2^{j_1}\delta$  and  $\varepsilon_2 = \sigma\alpha_1^{i_2}\gamma\alpha_2^{j_2}\delta$  be two words in  $S$ . Obviously, these words are incomparable if and only if either  $i_1 < i_2$  and  $j_1 > j_2$  or  $i_1 > i_2$  and  $j_1 < j_2$ . This implies that there are at most  $i_1 + j_1$  words in  $S$  incomparable with  $\varepsilon_1$ . Since for each word in  $S$  there exists only finitely many incomparable words, the set  $S$  does not contain an infinite antichain.  $\square$

For any set of words  $S$  we denote  $\langle S \rangle$  to be the set of factors of words in  $S$ . With this notation we extend the result of the previous lemma to the following.

**Lemma 2.2.6.** *Let  $\alpha_1, \alpha_2, \gamma$  be words. Then the set  $\langle \alpha_1^i\gamma\alpha_2^j \mid i, j \in \mathbb{N} \rangle$  is well-quasi-ordered.*

*Proof.* Let  $A_{1,p}, G_p, A_{2,p}$  be the sets of prefixes of  $\alpha_1, \gamma, \alpha_2$  and the corresponding sets of suffixes be  $A_{1,s}, G_s, A_{2,s}$ . These sets are clearly finite. Then  $\langle \alpha_1^i\gamma\alpha_2^j \mid i, j \in \mathbb{N} \rangle = \bigcup_{\sigma \in A_{1,s}, \delta \in A_{2,p}} \{\sigma\alpha_1^i\gamma\alpha_2^j\delta \mid i, j \in \mathbb{N}\} \bigcup_{\gamma_s \in G_s, \delta \in A_{2,p}} \{\gamma_s\alpha_2^j\delta \mid j \in \mathbb{N}\} \bigcup_{\delta' \in A_{2,s}, \delta \in A_{2,p}}$

$\{\delta'\alpha_2^j\delta \mid j \in \mathbb{N}\} \cup_{\sigma \in A_{1,s}, \gamma_p \in G_p} \{\sigma\alpha_1^i\gamma_p \mid i \in \mathbb{N}\} \cup_{\sigma \in A_{1,s}, \sigma' \in A_{1,p}} \{\sigma\alpha_1^i\sigma' \mid i \in \mathbb{N}\}$  is a finite union of sets which are well-quasi-ordered by Lemma 2.2.5 and Lemma 2.2.2.  $\square$

## 2.3 Main results

In this section, we show how to decide for a given finite deterministic automaton  $A = (G, q_0, Q)$  whether the language  $L(A)$  contains an infinite antichain or not with respect to the factor containment relation. Our solution is based on the analysis of the structure of cycles in  $G$ .

A cycle in  $G$  is any directed path with at least one edge in which the first and the last nodes coincide. A cycle is *simple* if its nodes are pairwise distinct (except for the first node being equal to the last node). Given a simple cycle  $C$ , we denote by  $|C|$  the *length* of  $C$ , i.e. the number of nodes in  $C$ . For a node  $v$  of  $C$ , we denote by  $w(C, v)$  the word of length  $|C|$  obtained by reading the labels of the edges of  $C$  starting from the node  $v$ .

We distinguish between two basic cases: the case where  $G$  contains two different simple cycles that have at least one node in common and the case where all simple cycles of  $G$  are pairwise node disjoint.

**Proposition 2.3.1.** *Let  $A = (G, q_0, Q)$  be a reduced finite deterministic automaton. If  $G$  contains two different simple cycles which have a node in common, then the language  $L(A)$  contains an infinite antichain.*

*Proof.* Let  $C_1$  and  $C_2$  be two different simple cycles with a common node  $v$ . Since the cycles are different we may assume without loss of generality that the node of  $C_1$  following  $v$  is different from the node of  $C_2$  following  $v$ . As a result, the edges of  $C_1$  and  $C_2$  leaving  $v$  are labeled with different letters of the alphabet.

We denote  $\alpha = w(C_1, v)^{|C_2|}$  and  $\beta = w(C_2, v)^{|C_1|}$ . The words  $\alpha$  and  $\beta$  have the same length, but they differ in the first letter according to the above assumption.

Since  $A$  is reduced, there exist a directed path  $\rho$  from the start node to  $v$  and a directed path  $\pi$  from  $v$  to a terminal node. Therefore, every word of the form  $w(\rho)\beta\alpha^i\beta w(\pi)$  ( $i = 1, 2, \dots$ ) belongs to the language  $L(A)$ . Since the words  $\alpha$  and  $\beta$  have the same length and differ in the first letter, we conclude that  $\beta$ , and hence  $w(\rho)\beta$ , is not a left extension of a power of  $\alpha$ . Similarly, the word  $\beta w(\pi)$  is not a right extension of a power of  $\alpha$ . Therefore, by Lemma 2.2.3, the language  $L(A)$  contains an infinite antichain.  $\square$



From now on, we consider automata in which every two simple cycles are node disjoint. In this case, we decompose the set of nodes into finitely many subsets of simple structure, called metapaths.

A *metapath* consists of a number of node disjoint simple cycles, say  $C_1, \dots, C_t$  (possibly  $t = 0$ ), and a number of directed paths  $\rho_0, \dots, \rho_t$  such that  $\rho_0$  connects the start node  $q_0$  to  $C_1$ ,  $\rho_1$  connects  $C_1$  to  $C_2$ ,  $\rho_2$  connects  $C_2$  to  $C_3$ , and so on, and finally,  $\rho_t$  connects  $C_t$  to a terminal node of the automaton. Let us observe that  $\rho_0$  and  $\rho_t$  can be of length 0, while all the other paths are necessarily of length at least one, since the cycles are node disjoint.

We denote by  $s(\rho_i)$  and  $f(\rho_i)$  the first and the last node of  $\rho_i$ , respectively, and observe that for  $i > 0$ ,  $s(\rho_i)$  belongs to  $C_i$ , and for  $i < t$ ,  $f(\rho_i)$  belongs to  $C_{i+1}$ .

If  $t = 0$ , then the metapath contains no cycles and consists of the path  $\rho_0$  alone. This path connects the start node  $q_0$  to a terminal node of the automaton and no node of this path belongs to a simple cycle.

If  $t > 0$ , then  $f(\rho_0), s(\rho_1), f(\rho_1), \dots, s(\rho_{t-1}), f(\rho_{t-1}), s(\rho_t)$  are the only nodes of the paths  $\rho_0, \rho_1, \dots, \rho_t$  that belong to simple cycles.

For  $i = 1, \dots, t$ , we denote by

- $\pi_i$  the directed path from  $f(\rho_{i-1})$  to  $s(\rho_i)$  taken along the cycle  $C_i$ ,
- $\gamma_i$  the word  $w(\pi_i)w(\rho_i)$ ,
- $\alpha_i$  the word  $w(C_i, f(\rho_{i-1}))$ .

Also, by  $\gamma_0$  we denote the word  $w(\rho_0)$ .

Let  $\tau$  be a metapath with  $t$  cycles, as defined above. The set of words accepted by this metapath can be described as follows: if  $t = 0$  then  $L(\tau) = \{\gamma_0\}$ , and if  $t > 0$  then

$$L(\tau) = \{\gamma_0 \alpha_1^{j_1} \gamma_1 \dots \gamma_{t-1} \alpha_t^{j_t} \gamma_t : j_1, \dots, j_t = 0, 1, \dots\}.$$

Clearly, the set  $T(A)$  of all metapaths is finite and

$$L(A) = \bigcup_{\tau \in T(A)} L(\tau).$$

It is also clear that  $L(A)$  contains an infinite antichain if and only if  $L(\tau)$  contains an infinite antichain for at least one metapath  $\tau \in T(A)$ .

If  $t = 0$ , the set  $L(\tau)$  is finite and hence cannot contain an infinite antichain. In order to determine if  $L(\tau)$  contains an infinite antichain for  $t > 0$ , we distinguish

between the following three cases:  $t = 1$ ,  $t = 2$  and  $t \geq 3$ . In our analysis below we use the following simple observation:

**Observation 2.3.1.** *For  $i = 1, \dots, t - 1$ , the word  $\gamma_i$  is not a right extension of a power of  $\alpha_i$ .*

The validity of this observation is due to the fact that the edge of  $\rho_i$  and the edge of  $C_i$  leaving vertex  $s(\rho_i)$  must have different labels. On the other hand, we note that  $\gamma_0$  may be a left extension of a power of  $\alpha_1$ , while  $\gamma_t$  may be a right extension of a power of  $\alpha_t$ .

**Proposition 2.3.2.** *Let  $\tau$  be a metapath with exactly one cycle. Then  $L(\tau)$  contains an infinite antichain if and only if*

- *neither  $\gamma_0$  is a left extension of a power of  $\alpha_1$*
- *nor  $\gamma_1$  is a right extension of a power of  $\alpha_1$ .*

*Proof.* If  $\gamma_0$  is a left extension of a power of  $\alpha_1$  or  $\gamma_1$  is a right extension of a power of  $\alpha_1$ , then  $L(\tau)$  does not contain an infinite antichain by Lemma 2.2.2.

If neither  $\gamma_0$  is a left extension of a power of  $\alpha_1$  nor  $\gamma_1$  is a right extension of a power of  $\alpha_1$ , then  $L(\tau)$  contains an infinite antichain by Lemma 2.2.3.  $\square$

**Proposition 2.3.3.** *Let  $\tau$  be a metapath with exactly two cycles. Then  $L(\tau)$  contains an infinite antichain if and only if*

- *either  $\gamma_0$  is not a left extension of a power of  $\alpha_1$*
- *or  $\gamma_2$  is not a right extension of a power of  $\alpha_2$ .*

*Proof.* Assume  $\gamma_0$  is not a left extension of a power of  $\alpha_1$ . From Observation 2.3.1 we know that  $\gamma_1$  is not a right extension of a power of  $\alpha_1$ . This implies that  $\gamma_1\gamma_2$  is not a right extension of a power of  $\alpha_1$ . Therefore, by Lemma 2.2.3, the set  $\{\gamma_0\alpha_1^i\gamma_1\gamma_2 : i = 1, 2, \dots\}$  is an infinite antichain. Since this set is a subset of  $L(\tau)$ , we conclude that  $L(\tau)$  contains an infinite antichain.

Suppose now that  $\gamma_2$  is not a right extension of a power of  $\alpha_2$ . By Observation 2.3.1,  $\gamma_1$  is not a right extension of a power of  $\alpha_1$ . Therefore, by Lemma 2.2.4,  $\gamma_0\alpha_1^{|\alpha_2|}\gamma_1$  is not a left extension of a power of  $\alpha_2$ . This implies, by Lemma 2.2.3, that the set  $\{\gamma_0\alpha_1^{|\alpha_2|}\gamma_1\alpha_2^i\gamma_2 : i = 1, 2, \dots\}$  is an infinite antichain. Since this set is a subset of  $L(\tau)$ , we conclude that  $L(\tau)$  contains an infinite antichain.

Finally, assume that  $\gamma_0$  is a left extension of a power of  $\alpha_1$  and  $\gamma_2$  is a right extension of a power of  $\alpha_2$ . Then the set  $L(\tau) = \{\gamma_0\alpha_1^i\gamma_1\alpha_2^j\gamma_2 \mid i, j \in \mathbb{N}\}$  satisfies the conditions of Lemma 2.2.5, and therefore is well-quasi-ordered.  $\square$

**Proposition 2.3.4.** *Let  $\tau$  be a metapath with  $t \geq 3$  cycles. Then  $L(\tau)$  contains an infinite antichain.*

*Proof.* By Observation 2.3.1,  $\gamma_1$  is not a right extension of a power of  $\alpha_1$ , and hence, by Lemma 2.2.4,  $\gamma_0\alpha_1^{|\alpha_2|}\gamma_1$  is not a left extension of a power of  $\alpha_2$ . Also, by Observation 2.3.1,  $\gamma_2$  is not a right extension of a power of  $\alpha_2$ , and hence  $\gamma_2 \dots \gamma_t$  is not a right extension of a power of  $\alpha_2$ . Therefore, by Lemma 2.2.3, the set  $\{\gamma_0\alpha_1^{|\alpha_2|}\gamma_1\alpha_2^i\gamma_2 \dots \gamma_t : i = 0, 1, 2, \dots\}$  is an infinite antichain. Since this set is a subset of  $L(\tau)$ , we conclude that  $L(\tau)$  contains an infinite antichain.  $\square$

We summarize the above discussion in the following final statement which is the main result of this chapter.

**Theorem 2.3.1.** *Given a deterministic finite automaton  $A$  one can decide in polynomial time whether the language accepted by  $A$  contains an infinite antichain with respect to the factor containment relation or not.*

*Proof.* Since the automaton  $A$  is finite, the question of the existence of infinite antichains in  $L(A)$  is decidable by Propositions 2.3.1, 2.3.2, 2.3.3, 2.3.4. Now we sketch the proof of polynomial-time solvability.

First, we identify strongly connected components in  $G$ , which can be done in polynomial time. If at least one strongly connected component contains a simple cycle and is different from the cycle (i.e. contains at least one edge outside of the cycle), then it necessarily contains two cycles with a common node, in which case  $L(A)$  contains an infinite antichain by Proposition 2.3.1.

If each of the strongly connected components of  $G$  is a simple cycle (or a single vertex without loops), then the cycles of  $G$  are pairwise node disjoint. In this case, we construct an auxiliary acyclic graph  $G'$  by contracting each simple cycle into a single node, called *cyclic node*. Each metapath in  $G$  corresponds to a directed path in  $G'$ .

First, we check if  $G'$  contains a directed path containing at least 3 cyclic nodes. This can be done for  $G'$  in cubic time. If such a path exists, then by Proposition 2.3.4  $L(A)$  necessarily contains an infinite antichain.

In order to check if  $G$  contains a metapath with precisely two cycles and satisfying conditions of Proposition 2.3.3, we choose in  $G'$  an ordered pair of cyclic nodes  $c_1, c_2$  and delete from the graph all other cyclic nodes, obtaining in this way the graph  $G''$ . The nodes  $c_1$  and  $c_2$  correspond to the cycles  $C_1$  and  $C_2$  in  $G$ . We assume that in  $G''$  there is a directed path  $\rho_0$  connecting  $q_0$  to  $c_1$ , a directed path  $\rho_1$  connecting  $c_1$  to  $c_2$  and a directed path  $\rho_2$  connecting  $c_2$  to a terminal node. Any

three such paths must be node disjoint (except, of course,  $c_1$  and  $c_2$ ), since otherwise a directed cycle arises. Therefore, together with  $C_1$  and  $C_2$  any three such paths form a metapath in  $G$ . Our task is to verify if there is a triple  $(\rho_0, \rho_1, \rho_2)$  that defines a metapath satisfying conditions of Proposition 2.3.3.

Let us observe that every choice of  $\rho_0$  uniquely defines the word  $\alpha_1$  (inscribed in  $C_1$ ) which we denote by  $\alpha_1^{\rho_0}$ . Therefore, in order to verify the first of the two conditions of Proposition 2.3.3 we need to solve the following problem.

- (1) Determine if  $G''$  contains a path  $\rho_0$  connecting  $q_0$  to  $c_1$  such that the corresponding word  $\gamma_0 = w(\rho_0)$  is not a left extension of a power of  $\alpha_1^{\rho_0}$ .

The second of the two conditions of Proposition 2.3.3 involves the word  $\alpha_2$ , which is defined by the path  $\rho_1$ . However, it turns out that this condition can be verified without specifying this path. In order to show this, let us associate with every path  $\rho_2$  connecting  $c_2$  to a terminal node the word  $\alpha_2^{\rho_2}$  defined as  $w(C_2, s(\rho_2))$ , i.e.  $\alpha_2^{\rho_2}$  is the word of length  $|C_2|$  inscribed in  $C_2$  starting at the first node of  $\rho_2$ . Then  $\gamma_2 = w(\pi_2)w(\rho_2)$  is not a right extension of a power of  $\alpha_2$  if and only if  $w(\rho_2)$  is not a right extension of a power of  $\alpha_2^{\rho_2}$ . Therefore, in order to verify the second of the two conditions of Proposition 2.3.3 we need to solve the following problem.

- (2) Determine if  $G''$  contains a path  $\rho_2$  connecting  $c_2$  to a terminal node such that  $w(\rho_2)$  is not a right extension of a power of  $\alpha_2^{\rho_2}$ .

To solve problem (1), we denote by  $n$  the number of nodes of  $G''$  and by  $N_i$  the set of nodes of  $G''$  for which there exists a directed path of length  $i$  from  $q_0$  ( $N_0 = \{q_0\}$ ). We observe that the sets  $N_i$  are not necessarily disjoint, each of them contains at most  $n$  nodes and for  $i > n$ , the sets  $N_i$  are empty (since  $G''$  is acyclic). Consider a vertex  $x \in N_i$  with  $i > 0$  and assume there is a directed path connecting  $x$  to  $c_1 \in N_j$  with  $j > i$  (which can be easily verified). We check if at least two edges coming to  $x$  from the vertices of  $N_{i-1}$  are labelled with different letters of the alphabet. If this is the case, then clearly there is a path  $\rho_0$  connecting  $q_0$  to  $c_1$  through  $x$  such that the corresponding word  $\gamma_0 = w(\rho_0)$  is not a left extension of a power of  $\alpha_1^{\rho_0}$ . If for each  $i$  and for each vertex  $x \in N_i$  all edges coming to  $x$  from  $N_{i-1}$  have the same label and this label coincides with the respective letter of  $\alpha_1^{\rho_0}$  (counted cyclically from the end of  $\alpha_1^{\rho_0}$ ), then problem (1) has no positive solution. It is not difficult to see that the overall time complexity of problem (1) is polynomial in the number vertices of  $G$ .

To solve problem (2), for each vertex  $v$  of  $C_2$  we check if  $v$  can be connected to a terminal node by at least two different paths. If this is the case, then at least

one of them is a path  $\rho_2$  such that the word  $w(\rho_2)$  is not a right extension of a power of  $\alpha_2^{\rho_2}$ , because the edges leaving the node where the two paths split must be labeled differently. If for each  $v \in C_2$  there is at most one path  $\rho_2$  connecting  $v$  to a terminal node and the word  $w(\rho_2)$  is a right extension of a power of  $\alpha_2^{\rho_2}$ , then problem (2) has no positive solution. Clearly, problem (2) can be solved in polynomial time too.

If at least one of the two problems, say (1), has a positive solution, then, in addition to this solution, we find an arbitrary directed path connecting  $c_1$  to  $c_2$  and an arbitrary directed path connecting  $c_2$  to a terminal node. Together with the cycles  $C_1$  and  $C_2$  these three paths form a metapath  $\tau$  in  $G$  such that  $L(\tau)$  contains an infinite antichain by Proposition 2.3.3.

The above arguments show that in polynomial time we can check if  $G$  has a metapath  $\tau$  with precisely two cycles such that  $L(\tau)$  contains an infinite antichain. If this is not the case, we need to check if  $G$  contains a metapath with exactly one cycle satisfying conditions of Proposition 2.3.2. This can be done by solving for each cyclic node of  $G'$  two problems similar to problems (1) and (2).  $\square$

## 2.4 An alternative approach

In the attempt to extend the solution for languages to other combinatorial structures (permutations, graphs, etc.), in this section we propose an alternative approach to the same problem, which is not based on the notion of automaton. We discuss possible ways to apply this approach to graphs and permutations in the next section.

The alternative approach is based on the notion of a *periodic* infinite antichain, which can be defined as follows.

**Definition 2.4.1.** An infinite antichain of words is periodic of period  $p$  if it has the form  $\{\beta\alpha^k\gamma \mid k \in \mathbb{N}\}$ , where  $|\alpha| = p$  and neither  $\beta$  is a right extension of a power of  $\alpha$  nor  $\gamma$  is a left extension of a power of  $\alpha$ .

Notice that Lemma 2.2.3 verifies that the set  $\{\beta\alpha^k\gamma \mid k \in \mathbb{N}\}$  is an antichain indeed. The notion of periodic infinite antichains and the results of Section 2.3 allow us to derive the following criterion of well-quasi-orderability of factorial languages defined by finitely many forbidden factors.

**Theorem 2.4.1.** Let  $D = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$  be a finite set of words and  $X = \text{Free}(\alpha_1, \dots, \alpha_k)$  the factorial language with the antidictionary  $D$ . Then  $X$  is well-quasi-ordered by the factor containment relation if and only if it contains no periodic infinite antichains of period at most  $(|\alpha_1| + |\alpha_2| + \dots + |\alpha_k| + 1)^2$ .

*Proof.* If  $X$  is well-quasi-ordered, then it certainly does not contain any periodic antichains.

Conversely, suppose  $X$  is not well-quasi-ordered. From the construction of the automaton for factorial languages given by Crochemore et al. [1998] we know that the number of nodes in the automaton is precisely the number of different prefixes of the forbidden words. Hence the size of the automaton is at most  $t = |\alpha_1| + |\alpha_2| + \dots + |\alpha_k| + 1$ .

Now, if the automaton contains two cycles  $C_1$  and  $C_2$  intersecting at some vertex  $v$ , Proposition 2.3.1 shows that the automaton contains an infinite antichain  $\{\beta\alpha^i\beta \mid i \in \mathbb{N}\}$ , where  $\alpha = w(C_1, v)^{|C_2|}$  and  $\beta = w(C_2, v)^{|C_1|}$ . Hence  $X$  contains a periodic infinite antichain of period  $|\alpha| = |C_1||C_2| \leq t^2$ .

Suppose now the automaton contains a metapath with at least three cycles. Then by Proposition 2.3.4 the set  $\{\gamma_0\alpha_1^{|\alpha_2|}\gamma_1\alpha_2^i\gamma_2 \mid i \in \mathbb{N}\}$  is a periodic infinite antichain of period  $|\alpha_1| \leq t$ .

Consider now the case when the automaton contains neither two intersecting cycles, nor a metapath with at least three cycles. Then, as  $X$  is not well-quasi-ordered, from Propositions 2.3.2 and 2.3.3 we conclude that  $X$  contains one of the following antichains  $\{\gamma_0\alpha_1^i\gamma_1 \mid i \in \mathbb{N}\}$  or  $\{\gamma_0\alpha_1^i\gamma_1\gamma_2 \mid i \in \mathbb{N}\}$  or  $\{\gamma_0\alpha_1^{|\alpha_2|}\gamma_1\alpha_2^i\gamma_2 \mid i \in \mathbb{N}\}$ . These are again periodic infinite antichains of period either  $|\alpha_1| \leq t$  or  $|\alpha_2| \leq t$ . Hence we conclude that if  $X$  is not well-quasi-ordered, then it contains a periodic antichain of period at most  $t^2$ . This completes the proof.  $\square$

The importance of this result is due to the fact that it suggests possible ways to approach the question of well-quasi-orderability for other combinatorial structures, such as graphs or permutations. We discuss this idea in Section 2.6.

Theorem 2.4.1 also raises an interesting question of determining the minimum value of the period in periodic infinite antichains that have to be broken to ensure well-quasi-orderability. According to the theorem, this value is at most  $(|\alpha_1| + |\alpha_2| + \dots + |\alpha_k| + 1)^2$ . We believe that this bound can be substantially improved and discuss this question in the next section for factorial languages with binary alphabet.

## 2.5 Deciding WQO for factorial languages with binary alphabet

Throughout this section we deal with words in the binary alphabet  $A = \{0, 1\}$ . If a word  $\beta$  is not a left extension of a power of a word  $\alpha$ , then we say that  $\beta$  is a minimal word with this property if any proper suffix of  $\beta$  is a left extension of a

power of  $\alpha$ . Similarly, we define the notion of a minimal word which is not a right extension of a power  $\alpha$ . It is not difficult to see that if  $\beta$  is a minimal not left/right extension of a power of  $\alpha$ , then  $|\beta| \leq |\alpha|$ .

The notions of minimal not left/right extensions allow us to define the notion of a minimal periodic infinite antichain as follows.

**Definition 2.5.1.** A periodic infinite antichain  $\{\beta\alpha^k\gamma \mid k \in \mathbb{N}\}$  is minimal if  $\beta$  is a minimal not left extension of a power of  $\alpha$  and  $\gamma$  is a minimal not right extension of a power of  $\alpha$ .

To illustrate this notion we list all minimal periodic infinite antichains of period at most 3 in Tables 2.1 and 2.2.

	Period 1		Period 2
(1.1)	0 1111 ... 1111 0	(2.1)	00 1010 ... 0101 00
(1.2)	1 0000 ... 0000 1	(2.2)	11 0101 ... 0101 00
		(2.3)	00 1010 ... 1010 11
		(2.4)	11 0101 ... 1010 11

Table 2.1: Minimal periodic infinite antichains of period 1 and 2

(3.1)	111 011011 ... 110110 111	(3.10)	000 100100 ... 001001 000
(3.2)	010 110110 ... 110110 111	(3.11)	101 001001 ... 001001 000
(3.3)	00 110110 ... 110110 111	(3.12)	11 001001 ... 001001 000
(3.4)	111 011011 ... 011011 010	(3.13)	000 100100 ... 100100 101
(3.5)	010 110110 ... 011011 010	(3.14)	101 001001 ... 100100 101
(3.6)	00 110110 ... 011011 010	(3.15)	11 001001 ... 100100 101
(3.7)	111 011011 ... 011011 00	(3.16)	000 100100 ... 100100 11
(3.8)	010 110110 ... 011011 00	(3.17)	101 001001 ... 100100 11
(3.9)	00 110110 ... 011011 00	(3.18)	11 001001 ... 100100 11

Table 2.2: Minimal periodic infinite antichains of period 3

We will say that a forbidden word  $\alpha$  destroys an infinite antichain  $A$  if  $\alpha$  is a factor of all members of  $A$ , except possibly finitely many of them. Theorem 2.4.1 tells us that in order to ensure well-quasi-orderability of a factorial language  $Free(\alpha_1, \dots, \alpha_k)$  it suffices to destroy periodic infinite antichains of period at most  $(|\alpha_1| + |\alpha_2| + \dots + |\alpha_k| + 1)^2$ . We believe that in case of binary words we can do much better and conjecture the following.

**Conjecture 2.5.1.** Let  $Z$  be a finite set of words. Then the language  $Free(Z)$  is well-quasi-ordered by the factor containment relation if and only if it does not contain minimal periodic infinite antichains of period at most  $|Z|$ .

Below we verify this conjecture for factorial languages with at most 3 forbidden words.

**Lemma 2.5.1.** *Let  $\alpha$  be a binary word. Then  $\text{Free}(\alpha)$  is well-quasi-ordered by the factor containment relation if and only if it does not contain minimal periodic infinite antichains of period 1.*

*Proof.* If the class  $X = \text{Free}(\alpha)$  contains a periodic infinite antichain of period 1, then it is not well-quasi-ordered. So assume  $X$  does not contain minimal periodic infinite antichains of period 1. As  $X$  does not contain the antichain (1.1),  $\alpha$  must belong to the set  $E_{1.1} = \{01^k, 1^k, 1^k 0 \mid k \in \mathbb{N}\}$ . Similarly, as  $X$  does not contain (1.2),  $\alpha$  must belong to the set  $E_{1.2} = \{10^k, 0^k, 0^k 1 \mid k \in \mathbb{N}\}$ . Therefore  $\alpha \in E_{1.1} \cap E_{1.2} = \{0, 1, 01, 10\}$ .

Clearly, the sets  $\text{Free}(1) = \{0^k \mid k \in \mathbb{N}\}$  and  $\text{Free}(0) = \{1^k \mid k \in \mathbb{N}\}$  are well-quasi-ordered. The sets  $\text{Free}(01) = \{1^k 0^l \mid k, l \in \mathbb{N}\}$  and  $\text{Free}(10) = \{0^k 1^l \mid k, l \in \mathbb{N}\}$  are well-quasi-ordered by Lemma 2.2.5.  $\square$

**Lemma 2.5.2.** *Let  $\alpha_1, \alpha_2$  be binary words. Then  $\text{Free}(\alpha_1, \alpha_2)$  is well-quasi-ordered by the factor containment relation if and only if it does not contain minimal periodic infinite antichains of period at most 2.*

*Proof.* If the class  $X = \text{Free}(\alpha_1, \alpha_2)$  contains a periodic antichain of period at most 2 then it is not well-quasi-ordered. So assume  $X$  does not contain a periodic antichain of period at most 2. As  $X$  does not contain (1.1), one of  $\alpha_1$  and  $\alpha_2$  must belong to the set  $E_{1.1} = \{01^k, 1^k, 1^k 0 \mid k \in \mathbb{N}\}$ . Similarly, as  $X$  does not contain an infinite antichain (1.2), one of  $\alpha_1$  and  $\alpha_2$  must belong to the set  $E_{1.2} = \{10^k, 0^k, 0^k 1 \mid k \in \mathbb{N}\}$ . Now either one of them belongs to  $E_{1.1} \cap E_{1.2} = \{0, 1, 01, 10\}$ , or without loss of generality we may assume that  $\alpha_1 \in E_{1.1} \setminus E_{1.2} = \{01^k, 1^k, 1^k 0 \mid k \geq 2\}$  and  $\alpha_2 \in E_{1.2} \setminus E_{1.1} = \{10^k, 0^k, 0^k 1 \mid k \geq 2\}$ . In this latter case,  $\alpha_2$  has two consecutive 0's, so it does not destroy (2.4). Therefore,  $\alpha_1$  has to destroy (2.4) and in particular  $\alpha_1$  cannot contain three consecutive 1's, giving that  $\alpha_1 \in \{11, 110, 011\}$ . Similarly, we get that  $\alpha_2 \in \{00, 100, 001\}$ . We now notice that  $\text{Free}(011, 001)$  contains (2.2) and  $\text{Free}(110, 100)$  contains (2.3). Hence  $\text{Free}(\alpha_1, \alpha_2)$  does not contain a periodic antichain of period at most 2 if one of the forbidden words is  $\{0, 1, 01, 10\}$  or

$$(\alpha_1, \alpha_2) \in \{(11, 00), (11, 100), (11, 001), (110, 00), (110, 001), (011, 00), (011, 100)\}.$$

If one of the forbidden words is  $\{0, 1, 01, 10\}$ , then  $X$  is a subclass of  $\text{Free}(0)$ ,  $\text{Free}(1)$ ,  $\text{Free}(01)$  or  $\text{Free}(10)$ , hence well-quasi-ordered by Lemma 2.5.1. The classes  $\text{Free}(110, 001) = \langle (10)^k 1^l, (01)^k 0^l \mid k, l \in \mathbb{N} \rangle$  and  $\text{Free}(011, 100) = \langle 1^k (10)^l,$



$0^k(10)^l \mid k, l \in \mathbb{N}$  are well-quasi-ordered by Lemma 2.2.6. The rest 5 classes not containing periodic antichains of period 2 are subclasses of either  $Free(110, 001)$  or  $Free(011, 100)$ , hence well-quasi-ordered.  $\square$

**Lemma 2.5.3.** *Let  $\alpha_1, \alpha_2, \alpha_3$  be binary words. Then  $Free(\alpha_1, \alpha_2, \alpha_3)$  is well-quasi-ordered by the factor containment relation if and only if it does not contain minimal periodic infinite antichains of period at most 3.*

*Proof.* If the class  $X = Free(\alpha_1, \alpha_2, \alpha_3)$  contains a periodic antichain of period at most 3 then it is not well-quasi-ordered. So assume  $X$  does not contain a periodic antichain of period at most 3. As  $X$  does not contain (1.1), one of the three forbidden words must belong to the set  $E_{1.1} = \{01^k, 1^k, 1^k0 \mid k \in \mathbb{N}\}$ . Similarly, as  $X$  does not contain an infinite antichain (1.2), one of the three forbidden words must belong to the set  $E_{1.2} = \{10^k, 0^k, 0^k1 \mid k \in \mathbb{N}\}$ . So either one of the words belong to  $E_{1.1} \cap E_{1.2} = \{0, 1, 01, 10\}$  or without loss of generality we may assume that  $\alpha_1 \in E_{1.1} \setminus E_{1.2} = \{01^n, 1^n, 1^n0 \mid n \geq 2\}$  and  $\alpha_2 \in E_{1.2} \setminus E_{1.1} = \{10^m, 0^m, 0^m1 \mid m \geq 2\}$ , and hence the cases to analyze can be distributed into 9 groups:

(A)	$(\alpha_1, \alpha_2) \in \{(01^n, 0^m) \mid n, m \geq 2\}$	(E)	$(\alpha_1, \alpha_2) \in \{(1^n, 10^m) \mid n, m \geq 2\}$
(B)	$(\alpha_1, \alpha_2) \in \{(01^n, 10^m) \mid n, m \geq 2\}$	(F)	$(\alpha_1, \alpha_2) \in \{(1^n, 0^m1) \mid n, m \geq 2\}$
(C)	$(\alpha_1, \alpha_2) \in \{(01^n, 0^m1) \mid n, m \geq 2\}$	(G)	$(\alpha_1, \alpha_2) \in \{(1^n0, 0^m) \mid n, m \geq 2\}$
(D)	$(\alpha_1, \alpha_2) \in \{(1^n, 0^m) \mid n, m \geq 2\}$	(H)	$(\alpha_1, \alpha_2) \in \{(1^n0, 10^m) \mid n, m \geq 2\}$
		(I)	$(\alpha_1, \alpha_2) \in \{(1^n0, 0^m1) \mid n, m \geq 2\}$

Notice that the words in (E) can be obtained from (A) by complementation (swapping 0's with 1's), the words in (F) from (A) by complementation followed by reversion (changing  $w = v_1v_2 \dots v_n$  to  $\bar{w} = v_nv_{n-1} \dots v_1$ ), the words in (G) from (A) by reversion, the words in (H) from (C) by complementation and the words in (I) from (B) by reversion. Notice that if a tripple  $(\alpha_1, \alpha_2, \alpha_3)$  destroys all minimal antichains of period at most 3, then so does the complement and reversion of a tripple. Therefore, in finding all such triples we can restrict our attention to cases (A), (B), (C), (D).

If  $n = 2$  and  $m = 2$ , the words in (A), (B), (D) destroy all antichains, and we consider the class (C) containing an antichain (2.2). In this class  $\alpha_3$  has to destroy (2.2), and hence  $\alpha_3 \in \langle 11(01)^p, (01)^p00 \mid p \in \mathbb{N} \rangle$ . So we get two tripples  $(011, 100, 11(01)^p)$  and  $(011, 100, (01)^p00)$  that destroy all antichains.

If  $n \geq 3, m \geq 3$ , then  $Free(\alpha_1, \alpha_2)$  contains (2.1), (2.4), (3.9), (3.18). As  $\alpha_3$  destroys (2.1) and (2.4),  $\alpha_3$  must have zeroes and ones alternating. Now, as  $\alpha_3$  destroys (3.9), the number of 0's in  $\alpha_3$  is at most 1 and to destroy (3.18), the number of 1's in  $\alpha_3$  must be at most one. We get that  $\alpha_3 \in \{0, 1, 01, 10\}$ . Now we

will consider the remaining cases (A), (B), (C), (D) for  $n = 2$  and  $m \geq 3$  or  $m = 2$  and  $n \geq 3$  separately.

**Case (A)**  $(\alpha_1, \alpha_2) = (01^n, 0^m)$

If  $n = 2, m \geq 3$ ,  $Free(011, 0^m)$  contains the antichains (2.1), (2.2), (3.13)-(3.15). So  $\alpha_3$  destroys (2.1) and (2.2) and we get  $\alpha_3 \in \langle (01)^p 00 \mid p \in \mathbb{N} \rangle$ . But only  $\alpha_3 = 0100$  and  $\alpha_3 = 0101$  and subwords destroy (3.13)-(3.15). Also notice that when  $m > 3$ ,  $Free(011, 0^m)$  contains (3.10)-(3.12), and only  $\alpha_3 = 0100$  and subwords destroy these antichains. If  $n \geq 3, m = 2$ ,  $Free(01^n, 00)$  contains (2.4), (3.4), (3.5). So  $\alpha_3$  destroys (2.4) and we get  $\alpha_3 \in \langle 11(01)^p, (10)^p 11 \mid p \in \mathbb{N} \rangle$ . But only  $\alpha_3 = 11010$  and  $\alpha_3 = 1011$  and the subwords destroy (3.4) and (3.5). Also, when  $n > 3$ ,  $Free(01^n, 00)$  contains (3.1)-(3.3), and the maximal subword of 11010 destroying these antichains is 1101.

**Case (B)**  $(\alpha_1, \alpha_2) = (01^n, 10^m)$ .

If  $n = 2$  and  $m \geq 3$  then  $Free(011, 10^m)$  contains 2.1), (2.2), (3.13)-(3.15). By the case (A), to destroy these antichains,  $\alpha_3 = 0100$  or  $\alpha_3 = 0101$  or subwords of these. Also notice that  $\alpha_3 = 0101$  can be only considered for  $m = 3$  as  $Free(011, 10000, 0101)$  contains (3.10)-(3.12). The tripples for  $n \geq 3$  and  $m = 2$  follow from the previous paragraph by complementation.

**Case (C)**  $(\alpha_1, \alpha_2) = (01^n, 0^m 1)$

If  $n = 2$  and  $m \geq 3$ ,  $Free(011, 0^m 1)$  contains (2.1), (2.2) and (3.12). As  $\alpha_3$  destroys (2.1) and (2.2), we get  $\alpha_3 \in \langle (01)^p 00 \mid p \in \mathbb{N} \rangle$  and only  $\alpha_3 = 0100$  or a subword destroys (3.12). The tripples for  $n \geq 3$  and  $m = 2$  follow from the previous paragraph by reversion and complementation.

**Case (D)**  $(\alpha_1, \alpha_2) = (1^n, 0^m)$ .

If  $n = 2$  and  $m \geq 3$ ,  $Free(11, 0^m)$  contains the antichains (2.1) and (3.14). As  $\alpha_3$  destroys (2.1), we get  $\alpha_3 \in \langle 00(10)^k, (01)^k 00 \mid k \in \mathbb{N} \rangle$ . From these only 00101, 10100 and subwords of these words destroy (3.14). If  $m > 3$ ,  $Free(11, 0^m)$  contains antichains (3.10)-(3.13). As  $\alpha_3$  destroys (3.10),  $\alpha_3$  cannot contain an occurrence of 101, so  $\alpha_3$  given by the previous case restricts to a subword of 0010 or 0100. The tripples for  $n \geq 3, m = 2$  follow from the previous paragraph by complementation.

Hence  $Free(\alpha_1, \alpha_2, \alpha_3)$  does not contain a minimal periodic antichain of length at most 3 if and only if one of the forbidden words belong to  $\{0, 1, 01, 10\}$ , two of the forbidden words belong to

$$\{(11, 00), (11, 100), (11, 001), (110, 00), (110, 001), (011, 00), (011, 100)\}.$$

or  $(\alpha_1, \alpha_2, \alpha_3)$  is one of the triples (or complement or reversion of the tripples) given in the table below.

Case	$(\alpha_1, \alpha_2, \alpha_3)$	$Free(\alpha_1, \alpha_2, \alpha_3)$
(A)	$(011, 0^m, 0100)$	$\langle 1^k 0^r (10)^l \mid k, l, r \in \mathbb{N}, r < m \rangle$
(A)	$(011, 000, 0101)$	$\langle 1^k (001)^l, 1^k 0 (100)^l \mid k, l \in \mathbb{N} \rangle$
(A)	$(01^n, 00, 1011)$	$\langle 01^r (01)^k, 1^k (01)^l \mid k, l, r \in \mathbb{N}, r < n \rangle$
(A)	$(0111, 00, 11010)$	$\langle 1^k (011)^l, (01)^k (011)^l \mid k, l \in \mathbb{N} \rangle$
(A)	$(01^n, 00, 1101)$	$\langle 1^k 0, (01)^k 1^r 0 \mid k, l, r \in \mathbb{N}, r < n \rangle$
(B)	$(011, 1000, 0101)$	$\langle 0^k (100)^l, 1^k 0 (100)^l, 1^k (001)^l \mid k, l \in \mathbb{N} \rangle$
(B)	$(011, 10^m, 0100)$	$\langle 0^k (10)^l, 1^k 0^r (10)^l \mid k, l, r \in \mathbb{N}, r < m \rangle$
(C)	$(011, 001, 11(01)^p)$	$\langle 1^k (01)^r 0^l, (01)^k 0^l \mid k, l, r \in \mathbb{N}, r < p \rangle$
(C)	$(011, 001, (01)^p 00)$	$\langle 1^k (01)^r 0^l, 1^k (01)^l \mid k, l, r \in \mathbb{N}, r < p \rangle$
(C)	$(011, 0^m 1, 0100)$	$\langle 1^k 0^l, 1^k 0^r (10)^l \mid k, l, r \in \mathbb{N}, r < m \rangle$
(D)	$(11, 000, 00101)$	$\langle (10)^k (100)^l \mid k, l \in \mathbb{N} \rangle$
(D)	$(11, 000, 10100)$	$\langle (100)^k (10)^l \mid k, l \in \mathbb{N} \rangle$
(D)	$(11, 0^m, 0010)$	$\langle (01)^k 0^r 1 \mid k, l, r \in \mathbb{N}, r \leq m \rangle$
(D)	$(11, 0^m, 0100)$	$\langle 10^r (10)^k \mid k, l, r \in \mathbb{N}, r \leq m \rangle$

Table 2.3: Tripples  $(\alpha_1, \alpha_2, \alpha_3)$  destroying antichains of period at most 3

Now we know that the classes obtained in the first two cases are well-quasi-ordered by Lemmas 2.5.1 and 2.5.2. The first tripple in the table gives a well-quasi-ordered class  $Free(011, 0^m, 0100)$ , because it consists of union of  $m$  sets  $\langle 1^k 0^r (10)^l \mid k, l \in \mathbb{N} \rangle$ ,  $r = 0, 1, \dots, (m - 1)$  each of which is well-quasi-ordered by Lemma 2.2.6. Similarly, all the classes in the table can be easily seen as finite unions of sets which are well-quasi-ordered by Lemma 2.2.6. Moreover, as complementation and reversion preserve antichains, we conclude that all classes with three forbidden words destroying minimal periodic antichains of period at most 3 are well-quasi-ordered. □

## 2.6 Conclusion

The alternative solution to the problem of deciding well-quasi-orderability of factorial languages proposed in Section 2.4 suggests a possible way to approach the same problem for graphs and permutations. Similarly to languages, this approach is based on the notion of *periodic* infinite antichains and consists in checking the presence of antichains of only bounded periodicity. Below we outline this approach for the induced subgraph relation on graphs and briefly discuss it for the pattern containment relation on permutations.

To define the notion of a periodic infinite antichain for graphs, we use the notion of letter graphs introduced in [Petkovšek, 2002] and slightly adapt it to our purposes. Originally, this notion was defined as follows.

Let  $A$  be a finite alphabet and  $S \subseteq A^2$  a binary relation on  $A$ . With each word  $\alpha = (\alpha_1\alpha_2\ldots\alpha_n)$  over  $A$  we associate the graph  $G_\alpha$  whose vertex set is  $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$  and two vertices  $\alpha_i$  and  $\alpha_j$  with  $i < j$  are adjacent if and only if  $(\alpha_i, \alpha_j) \in S$ . If  $A$  consists of  $k$  letters, the graph  $G_\alpha$  is called a  $k$ -letter graph. The importance of this notion for well-quasi-orderability is due to the fact that for each fixed  $k$  the set of all  $k$ -letter graphs is well-quasi-ordered by the induced subgraph relation [Petkovšek, 2002].

We now modify the notion of letter graphs by distinguishing between consecutive and nonconsecutive vertices of  $\alpha$ . For nonconsecutive vertices  $\alpha_i$  and  $\alpha_j$  with  $i < j$  the definition remains the same:  $\alpha_i$  and  $\alpha_j$  are adjacent if and only if  $(\alpha_i, \alpha_j) \in S$ . For consecutive vertices, we change the definition to the opposite:  $\alpha_i$  and  $\alpha_{i+1}$  are adjacent if and only if  $(\alpha_i, \alpha_{i+1}) \notin S$ . Let us denote the graph obtained in this way from the word  $\alpha$  by  $G_\alpha^*$ . For instance, if  $a$  is a letter of  $A$  and  $(a, a) \notin S$ , then the word  $aaaaa$  defines a path on 5 vertices.

The graph  $G_\alpha^*$  constructed from a periodic word  $\alpha$  will be called a periodic graph. The period of  $\alpha$  will be called the period of  $G_\alpha^*$ . In this way we define periodic graphs. Now, provided two periodic words  $\alpha$  and  $\beta$  are sufficiently long, it can be shown that if  $G_\alpha^*$  is an induced subgraph of  $G_\beta^*$  then any induced subgraph embedding must map consecutive vertices of  $G_\alpha^*$  into consecutive vertices of  $G_\beta^*$ , i.e. the induced subgraph relation of graphs corresponds to factor relation on words. We will describe this correspondence between words and graphs more rigorously in Chapter 5. This observation allows us to construct periodic antichains: we only need to break the periodicity on both ends of the graphs (words).

This could be done by changing the letters of the endpoints of the periodic graphs, i.e. constructing graphs corresponding to infinite periodic antichains of words. For example, the antichain  $\mathcal{C} = \{bab, baab, baaab, \ldots\}$  with binary relation  $S = \{(b, b)\}$  corresponds to the antichain of cycles  $C_3, C_4, \ldots$ . Also, one can obtain the antichain of complements of cycles from the same antichain  $\mathcal{C}$  with binary relation  $S = \{(a, a), (a, b), (b, a)\}$ .

In order to capture the notion of periodicity in the antichains of graphs one may want to avoid considering different possibilities of breaking periodicity by adding or removing some edges involving endpoints of periodic graphs. Instead, we could color the first and the last vertices of each graph differently from the intermediate vertices (see Figure 2.1 for an illustration). If we now strengthen

the induced subgraph relation by requiring that an embedding of one graph into another should respect the colors, then we convert the set of paths  $P_3, P_4, P_5 \dots$  into an infinite periodic antichain of period 1.



Figure 2.1: A colored path on  $P_5$

To further justify this restriction to colored (also known as labelled) infinite antichain, let us observe that in Chapter 3 we conjecture that for hereditary classes of graphs defined by *finitely many* forbidden induced subgraphs, the notion of well-quasi-orderability by induced subgraphs coincides with the notion of well-quasi-orderability by labelled induced subgraphs and verify this conjecture for all known examples of well-quasi-ordered classes of graphs. Also, notice that according to the conjecture of Pouzet [1972], in the case of labelled induced subgraphs one can be restricted to two different labels (colors).

For a finite collection  $C$  of graphs, let us denote by  $t(C)$  the total number of vertices of graphs in  $C$ . Suggested by the result on languages, we conjecture the following.

**Conjecture.** *There is a function  $f : \mathbb{N} \rightarrow \mathbb{N}$  such that the class  $X$  of graphs defined by a finite collection  $C$  of forbidden induced subgraphs is well-quasi-ordered by the induced subgraph relation if and only if  $X$  contains no periodic infinite antichains of period at most  $f(t(C))$ .*

To support the conjecture, let us notice that in the case of one forbidden induced subgraph  $G$  the class  $\text{Free}(G)$  is well-quasi-ordered if and only if it contains no periodic infinite antichains of period 1, in particular the set of cycles and their complements. Now if  $G$  contains a cycle  $C_3$ ,  $C_4$  or  $C_5$ , then  $\text{Free}(G)$  contains the antichain of all cycles of length greater than 5 and hence is not well-quasi-ordered. Similarly, if  $G$  contains the complements of  $C_3$ ,  $C_4$  or  $C_5$ , then it contains the complements of the cycles of length greater than 5 and hence is not well-quasi-ordered either. If  $G$  is free of  $C_3$ ,  $C_4$ ,  $C_5$  and their complements, then  $G$  is a path  $P_4$  on 4 vertices (or its induced subgraph), in which case  $\text{Free}(G)$  is well-quasi-ordered [Damaschke, 1990].

We also believe that a similar approach should also work for the pattern containment on permutations. Indeed, with each permutation  $\pi : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$  we can associate the permutation graph  $G_\pi$  on the vertex set  $\{1, 2, \dots, n\}$  in which two vertices  $i, j$  are adjacent if and only if  $(i - j)(\pi(i) - \pi(j)) < 0$ . Then the

pattern containment on permutations corresponds to the induced subgraph relation on the permutation graphs.

## Chapter 3

# Graphs

In this chapter we study our fundamental problem for graphs: given a finite set of graphs  $Z$ , determine whether the hereditary class  $\text{Free}(Z)$  is a well-quasi-order for the induced subgraph and the labelled induced subgraph relations. We conjecture that finitely defined classes are well-quasi-ordered by the induced subgraph relation if and only if they are well-quasi-ordered by the labelled induced subgraph relation. In Section 3.2 we verify this conjecture for all hereditary classes of graphs which are known to be well-quasi-ordered by induced subgraphs. In Section 3.3 we support the conjecture by a number of new results. We will further work and obtain a more general result in Section 5, showing that the conjecture is true for all hereditary classes with finite distinguishing number.

The next section contains basic definitions and tools that have been used in the proof of well-quasi-orderability by induced subgraphs and we extend some of the tools to the case of labelled graphs.

### 3.1 Induced Subgraph vs Labelled Induced Subgraph

For the induced subgraph relation, only few examples of graph classes which are well-quasi-ordered under this relation are known in the literature (see e.g. [Korpelainen and Lozin, 2011b,a]). Most of these examples deal with classes defined by finitely many forbidden induced subgraphs. For some of these classes even stronger results are available, namely, some of them are well-quasi-ordered under the *labelled* induced subgraph relation.

To define the notion of labelled induced subgraphs, let us consider an arbitrary quasi-order  $(W, \leq)$ . We say that  $G$  is a *labelled graph* if each vertex  $v$  of  $G$  is equipped with an element  $l(v) \in W$  (the label of  $v$ ). Given two labelled graphs  $G$

and  $H$ , we say that  $G$  is a labelled-induced subgraph of  $H$  if  $G$  is isomorphic to an induced subgraph of  $H$  and the isomorphism maps each vertex  $v$  of  $G$  to a vertex  $w$  of  $H$  with  $l(v) \leq l(w)$ . Finally, we say that a class of graphs  $X$  is well-quasi-ordered under the *labelled* induced subgraph relation if it contains no infinite antichains of labelled graphs whenever  $(W, \leq)$  is a *well*-quasi-order.

Clearly, well-quasi-orderability by labelled induced subgraphs implies well-quasi-orderability by induced subgraphs. However, the converse statement is generally not true. Linear forests (i.e. graphs in which every connected component is a path) give an example of a class of graphs which is well-quasi-ordered by induced subgraphs and not by *labelled* induced subgraphs. It is not difficult to see that there are infinitely many minimal forbidden induced subgraphs for this class (in particular, all chordless cycles are minimal forbidden induced subgraphs). In fact, Daligault et al. [2010] proved that no infinitely defined class is well-quasi-ordered by labelled subgraphs. On the other hand, many tools to prove well-quasi-ordering extend naturally to labelled subgraphs, which leads us to expect that the relations only differ for classes defined by infinitely many forbidden subgraphs.

**Conjecture 3.1.1.** A hereditary class  $X$  of graphs which is well-quasi-ordered by the induced subgraph relation is also well-quasi-ordered by the labelled induced subgraph relation if and only if  $X$  is defined by finitely many minimal forbidden induced subgraphs.

In the rest of this section, we revise basic tools that have been used in the proof of well-quasi-orderability by induced subgraphs and extend some of them to the case of labelled graphs.

### 3.1.1 Higman's lemma and Kruskal's theorem

For an arbitrary set  $M$ , denote by  $M^*$  the set of all finite sequences of elements of  $M$ . Any quasi-order  $\leq$  on  $M$  defines a quasi-order  $\preceq$  on  $M^*$  as follows:  $(a_1, \dots, a_m) \preceq (b_1, \dots, b_n)$  if and only if there is an order-preserving injection  $f : \{a_1, \dots, a_m\} \rightarrow \{b_1, \dots, b_n\}$  with  $a_i \leq f(a_i)$  for each  $i = 1, \dots, m$ . We call  $\preceq$  the subsequence relation. The celebrated Higman's lemma [Higman, 1952] states

**Lemma 3.1.1.** *If  $(M, \leq)$  is a WQO, then  $(M^*, \preceq)$  is a WQO.*

Kruskal [1960] extended this result to the set of finite trees partially ordered under homeomorphic embedding. The author proved his theorem under the additional assumption that the vertices of trees are equipped with labels from a well-quasi-ordered set. Thus, Kruskal's tree theorem restricted to paths becomes Higman's lemma.



### 3.1.2 Modular decomposition

Given a graph  $G = (V, E)$ , a subset of vertices  $U \subseteq V$  and a vertex  $x \in V$  outside  $U$ , we say that  $x$  *distinguishes*  $U$  if  $x$  has both a neighbour and a non-neighbour in  $U$ . A subset  $U \subseteq V$  is called a *module* of  $G$  if no vertex in  $V \setminus U$  distinguishes  $U$ . A module  $U$  is *nontrivial* if  $1 < |U| < |V|$ , otherwise it is *trivial*. A graph is called *prime* if it has only trivial modules.

An important property of maximal modules is that if  $G$  and the complement of  $G$  are both connected, then the maximal modules of  $G$  are pairwise disjoint. Moreover, from the above definition it follows that if  $U$  and  $W$  are maximal modules, then  $U$  is either complete or anticomplete to  $W$ . Therefore, by contracting each maximal module of  $G$  into a single vertex we obtain an induced subgraph  $G^0$  of  $G$  which is prime. Sometimes this graph is called the characteristic graph of  $G$  (alternatively, you can think of  $G$  as being obtained from  $G^0$  by substituting its vertices by maximal modules of  $G$ ). This property allows to recursively decompose the graph into connected components, co-components or maximal modules. This decomposition can be described by a rooted tree and is known in the literature under the name of *modular decomposition*. The importance of this decomposition for the study of well-quasi-orderability is due to the following theorem proved in [Korpelainen and Lozin, 2011a].

**Theorem 3.1.1.** *If the set of prime graphs in a hereditary class  $X$  is well-quasi-ordered by the labelled induced subgraph relation, then the class  $X$  is well-quasi-ordered by the induced subgraph relation.*

In Section 3.1.4, we will extend Theorem 3.1.1 to a stronger result stating that well-quasi-orderability of prime graphs in a hereditary class  $X$  by the labelled induced subgraph relation implies well-quasi-orderability of all graphs in  $X$  by the labelled induced subgraph relation. We also prove a similar result for canonical decomposition of bipartite graphs in Section 3.1.5. Before that, we derive a number of useful properties of minimal infinite antichains in Section 3.1.3.

### 3.1.3 Minimal infinite antichains

In this section, we assume that we are given an arbitrary quasi-order  $\leq$  on graphs, which is not a well-quasi-order, and derive a number of useful properties of minimal infinite antichains. We start by updating some terminology and notation from the induced subgraph relation to the more general context of quasi-order  $\leq$ .

If  $H \leq G$ , we will say that  $G$  contains  $H$  (or  $H$  is contained in  $G$ ), and if  $H \leq G$  and  $H \neq G$ , then we will say that  $G$  *properly* contains  $H$  and will denote

it as  $H < G$ . A downward closed set (with respect to  $\leq$ ) is any set  $X$  such that  $G \in X$  implies  $H \in X$  for every graph  $H \leq G$ . In particular, a hereditary class is a downward closed set with respect to the induced subgraph relation. For an antichain  $A$ , we will denote by

- $Free(A)$  the set of those graphs that do not contain any graph from  $A$ ,
- $Below(A)$  the set of those graphs that are properly contained in some graphs of  $A$ .

**Observation 3.1.1.** *For any antichain  $A$ ,*

$$Below(A) \subseteq Free(A).$$

*Proof.* If a graph  $H \in Below(A)$  does not belong to  $Free(A)$ , then, by definition, it must contain a graph  $G \in A$ . On the other hand, by definition, there is a graph  $G' \in A$  such that  $H < G'$ . But then  $G \leq H < G'$ , which contradicts the fact that  $A$  is an antichain.  $\square$

Among the various definitions of a minimal infinite antichain we use the following one proposed by Gustedt [1998].

**Definition 3.1.1.** An infinite antichain  $A$  is called minimal if there is no infinite antichain  $B$  such that  $Free(B) \subset Free(A)$ .

Gustedt also proved that if a quasi-order contains an infinite antichain, then it also contains a minimal one. We will use this fact with respect to downward closed sets of graphs. To this end, let us introduced the following definition.

**Definition 3.1.2.** An infinite antichain  $A$  belonging to a downward closed set  $X$  will be called  $X$ -minimal if there is no infinite antichain  $B \subseteq X$  such that  $Free(B) \subset Free(A)$ .

In this terminology, the result of Gustedt can be stated as follows.

**Lemma 3.1.2.** *If a downward closed set  $X$  contains an infinite antichain, then it contains an  $X$ -minimal infinite antichain.*

We now prove the following helpful properties of minimal antichains.

**Lemma 3.1.3.** *Let  $X$  be a downward closed set. If  $A$  is an  $X$ -minimal infinite antichain, then*

$$(1) \quad Free(A) \cap X = Below(A),$$

(2)  $Free(A) \cap X$  is well-quasi-ordered.

*Proof.* To prove (1), we observe that for each graph  $G \in Free(A) \cap X$ , there must exist a graph in  $A$  comparable with  $G$ , since otherwise  $A \cup \{G\}$  is an antichain in  $X$  such that  $Free(A \cup \{G\}) \subset Free(A)$ , contradicting the  $X$ -minimality of  $A$ . On the other hand,  $G$  does not contain any graph from  $A$ , since  $G \in Free(A)$ . Therefore, for each graph  $G \in Free(A) \cap X$ , there is a graph in  $A$  containing  $G$ . Obviously, this containment is proper, since otherwise  $G$  could not belong to  $Free(A)$ , i.e.  $G \in Below(A)$ .

To prove (2), assume to the contrary that  $Free(A) \cap X$  contains an infinite antichain  $B$ . By (1),  $B \subset Below(A)$ . If we denote by  $C$  the set of those graphs in  $A$  that do not contain any graph of  $B$ , then  $B \cup C$  is an infinite antichain in  $X$ . Moreover,  $Free(B \cup C) \subset Free(A)$ , which contradicts the minimality of  $A$ .  $\square$

### 3.1.4 Modular decomposition

Lemmas 3.1.2 and 3.1.3 allow us strengthening Theorem 3.1.1 as follows.

**Theorem 3.1.2.** *If the prime graphs in a hereditary class  $X$  are well-quasi-ordered by the labelled induced subgraph relation, then all graphs in  $X$  are well-quasi-ordered by the labelled induced subgraph relation.*

*Proof.* Suppose by contradiction that there is an infinite antichain  $A$  of labelled graphs in  $X$  with respect to the labelled induced subgraph relation. We assume that the vertices of graphs in  $A$  are labelled by the elements of a certain well-quasi-ordered set  $W$ . By Lemma 3.1.2 we may assume that  $A$  is an  $X$ -minimal infinite antichain. Then  $A$  must contain non-prime graphs, since the set of prime graphs is well-quasi-ordered. For each non-prime graph  $G \in A$ , we contract each maximal module  $M$  of  $G$  into a single vertex  $m$  (called module vertex) and assign to this vertex the (labelled) graph  $G[M]$ , obtaining in this way a prime graph  $G'$ . This operation transforms  $A$  into a sequence  $A'$  of prime labelled graphs from  $X$ . The vertices of these graphs are labelled either by the elements of  $W$  or by some *proper* induced subgraphs of graphs in  $A$ . By Lemma 3.1.3, the set of proper induced subgraphs of graphs in  $A$ , i.e.  $Below(A)$ , is well-quasi-ordered. Therefore, all graphs in  $A'$  are labelled by the elements of the well-quasi-ordered set  $W^* = W \cup Below(A)$  (we define the elements of  $W$  and  $Below(A)$  to be incomparable). Then, by assumption,  $A'$  contains no infinite antichains with respect to the labelled induced subgraph relation. In other words, there must exist two labelled graphs  $G'_1$  and  $G'_2$  in  $A'$  and an embedding  $\phi'$  of  $G'_1$  into  $G'_2$  as an induced subgraph such that  $\phi'$  respects the

labellings of the graphs. We denote by  $G_1$  the graph of  $A$  which was contracted to  $G'_1$  and by  $G_2$  the graph of  $A$  which was contracted to  $G'_2$ . The embedding  $\phi'$  maps module vertices of  $G'_1$  into module vertices of  $G'_2$ . This defines a corresponding mapping of modules of  $G_1$  into modules of  $G_2$ , and therefore, an embedding  $\phi$  of  $G_1$  into  $G_2$  as an induced subgraph such that  $\phi$  respects the labellings of  $G_1$  into  $G_2$ . This contradicts the fact that  $A$  is an antichain.  $\square$

### 3.1.5 Canonical decomposition of bipartite graphs

For any two disjoint bipartite graphs  $G_1 = (X_1, Y_1, E_1)$  and  $G_2 = (X_2, Y_2, E_2)$  we define the following three binary operations:

- the disjoint union  $\oplus$  is the operation that creates out of  $G_1$  and  $G_2$  the bipartite graph  $G = (X_1 \cup X_2, Y_1 \cup Y_2, E_1 \cup E_2)$ ,
- the join  $\otimes$  is the operation that creates out of  $G_1$  and  $G_2$  the bipartite graph which is the bipartite complement of the disjoint union of  $\tilde{G}_1$  and  $\tilde{G}_2$ ,
- the skew join  $\odot$  is the operation that creates out of  $G_1$  and  $G_2$  the bipartite graph  $G = (X_1 \cup X_2, Y_1 \cup Y_2, E_1 \cup E_2 \cup \{xy : x \in X_1, y \in Y_2\})$ .

These three operations define a decomposition scheme, known as *canonical decomposition*, which takes a bipartite graph  $G$  and whenever  $G$  has one of the following three forms  $G = G_1 \oplus G_2$ ,  $G = G_1 \otimes G_2$ , or  $G = G_1 \odot G_2$ , partitions it into  $G_1$  and  $G_2$ , and then the scheme applies to  $G_1$  and  $G_2$  recursively. Graphs that cannot be decomposed into smaller graphs under this scheme are called *canonically prime*. In [Ding, 1992], this scheme was used (implicitly) to show well-quasi-orderability of some classes of bipartite graphs with respect to the induced subgraph relation by reducing the problem to canonically prime graphs in the class. Now we extend this idea to the more general context of the labelled induced subgraph relation.

**Theorem 3.1.3.** *Let  $X$  be a hereditary class of bipartite graphs. If canonically prime graphs in  $X$  are well-quasi-ordered by the labelled induced subgraph relation, then all graphs in  $X$  are well-quasi-ordered by the labelled induced subgraph relation.*

*Proof.* The proof follows the same strategy as that of Theorem 3.1.2, so we reduce it to a sketch. By contradiction we assume that there is an infinite  $X$ -minimal antichain  $A$  of *labelled* graphs in  $X$  with respect to the labelled induced subgraph relation with vertex labels coming from a well-quasi-ordered set  $W$ . Every graph  $G$  in  $A$  which is not canonically prime is represented by a word  $G_1 \odot G_2$ , where

$\odot \in \{\oplus, \otimes, \oslash\}$ . This transforms  $A$  into a sequence of finite words in a well-quasi-ordered alphabet (because  $G_1$  and  $G_2$  belong to  $\text{Below}(A)$ ). By Higman's lemma [Higman, 1952], this sequence must contain a pair of comparable words, i.e. two words  $G_1 \odot G_2$  and  $G'_1 \odot G'_2$  such that  $G_i$  can be embedded into  $G'_i$  as an induced subgraph so that the labeling of their vertices is respected (we assume that different operations in the set  $\{\oplus, \otimes, \oslash\}$  are incomparable). But then  $G = G_1 \odot G_2$  can be embedded into  $G' = G'_1 \odot G'_2$  as an induced subgraph so that the labeling of their vertices is respected. This contradicts the fact that  $A$  is an antichain and hence proves the theorem.  $\square$

## 3.2 An overview of graph classes

In this section we give an overview of hereditary classes of graphs that have been shown to be well-quasi-ordered with respect to the induced subgraph relation.

### 3.2.1 $k$ -letter graphs

The notion of a  $k$ -letter graph was introduced by Petkovšek [2002] as follows.

**Definition 3.2.1.** A  $k$ -letter graph  $G$  is a graph defined by a finite word  $x_1x_2 \dots x_n$  on alphabet  $X$  of size  $k$  together with a subset  $S \subseteq X^2$  such that:

- $V(G) = \{x_1, x_2, \dots, x_n\}$
- $E(G) = \{x_ix_j : i \leq j \text{ and } (x_i, x_j) \in S\}$

For any fixed sets  $X$  and  $S \subseteq X^2$ , the subsequence relation on words corresponds precisely to the induced subgraph relation on  $k$ -letter graphs. Since there are only finitely many different choices for  $S$ , the following is an immediate corollary of Higman's lemma:

**Corollary 3.2.1.** [Petkovšek, 2002] *For any fixed  $k$ , the class of  $k$ -letter graphs is well-quasi-ordered by the induced subgraph relation.*

Using Higman's lemma in all its generality (which is just a special case of Kruskal's tree theorem), the above corollary can be extended in the following way.

**Theorem 3.2.1.** *For any fixed  $k$ , the class of  $k$ -letter graphs is well-quasi-ordered by the labelled induced subgraph relation.*

The fact that, for any fixed  $k$ , the class of  $k$ -letter graphs is described by finitely many forbidden induced subgraphs now readily follows from the result of

Daligault et al. [2010] concerning labelled well-quasi-ordered classes and was also proved explicitly in [Petkovšek, 2002]. Therefore, Conjecture 3.1.1 is valid for  $k$ -letter graphs.

### 3.2.2 $k$ -uniform graphs

Let  $k$  be a natural number,  $K$  a symmetric 0-1 square matrix of order  $k$ , and  $F_k$  a simple graph on the vertex set  $\{1, 2, \dots, k\}$ . Let  $H$  be the disjoint union of infinitely many copies of  $F_k$ , and for  $i = 1, \dots, k$ , let  $V_i$  be the subset of  $V(H)$  containing vertex  $i$  from each copy of  $F_k$ . Now we construct from  $H$  an infinite graph  $H(K)$  on the same vertex set by connecting two vertices  $u \in V_i$  and  $v \in V_j$  if and only if  $uv \in E(H)$  and  $K(i, j) = 0$  or  $uv \notin E(H)$  and  $K(i, j) = 1$ . Finally, let  $\mathcal{P}(K, F_k)$  be the hereditary class consisting of all the finite induced subgraphs of  $H(K)$ .

**Definition 3.2.2.** A graph  $G$  is called  $k$ -uniform if there is a number  $k$  such that  $G \in \mathcal{P}(K, F_k)$  for some  $K$  and  $F_k$ .

The following theorem was proved in [Korpelainen and Lozin, 2011a].

**Theorem 3.2.2.** *For any fixed  $k$ , the set of  $k$ -uniform graphs is well-quasi-ordered by the labelled induced subgraph relation.*

As before, we can deduce that these classes are finitely defined by the result of Daligault et al. [2010]. Alternatively, we can verify this directly as follows.

**Theorem 3.2.3.** *For any fixed  $k$ , the set of all  $k$ -uniform graphs is described by finitely many minimal forbidden induced subgraphs.*

*Proof.* In [Korpelainen and Lozin, 2011a], it was shown that if  $v$  is a vertex of a graph  $G$  and  $G - v$  is a  $k$ -uniform graph, then  $G$  is  $2k + 1$ -uniform. Therefore, all minimal non- $k$ -uniform graphs are  $2k + 1$ -uniform. Since the set of minimal non- $k$ -uniform graphs is necessarily an antichain, we conclude by Theorem 3.2.2 that this set is finite.  $\square$

This verifies Conjecture 3.1.1 for  $k$ -uniform graphs.

### 3.2.3 Monotone classes

A hereditary class of graphs is *monotone* if it is closed under taking subgraphs, not necessarily induced. [Ding, 1992] showed that a monotone class is well-quasi-ordered by the induced subgraphs relation if and only if it contains cycles  $C_n$  and graphs  $H_n$  for finitely many values of  $n$ . Moreover, he also showed that the class of graphs

containing no  $P_k$  as a subgraph (not necessarily induced) is well-quasi-ordered by the *labelled* induced subgraph relation.

Let  $X$  be a monotone class of graphs. It can be described either by the set  $M$  of minimal forbidden subgraphs or by the set  $M'$  of minimal forbidden *induced* subgraphs (the set  $M'$  can be obtained from  $M$  by adding to it all graphs containing graphs in  $M$  as spanning subgraphs). Clearly,  $M$  is finite if and only if  $M'$  is finite.

From the results of Guoli Ding it follows that if  $M$  is finite then  $X$  is well-quasi-ordered by induced subgraphs if and only if  $M$  contains a linear forest (otherwise  $X$  contains infinitely many cycles), in which case it is also well-quasi-ordered by labelled induced subgraphs. On the other hand, if  $X$  is well-quasi-ordered by induced subgraphs and the set  $M$  of its minimal forbidden subgraphs is infinite, then  $M$  cannot contain a linear forest (otherwise  $M$  contains non-minimal forbidden subgraphs), in which case  $X$  contains all linear forests and hence it is not well-quasi-ordered by the *labelled* induced subgraph relation. This discussion shows that Conjecture 3.1.1 is true in the family of monotone classes.

### 3.2.4 Other classes of graphs

Except for the three big families of graph classes mentioned earlier ( $k$ -letter graphs,  $k$ -uniform graphs and monotone classes), there are only finitely many other classes which are known to be well-quasi-ordered by induced subgraphs.

#### Subclasses of bipartite graphs

In this section, we list all maximal hereditary classes of bipartite graphs which are known to be well-quasi-ordered by induced subgraphs. Notice that in all these classes an induced path  $P_k$  (for some values of  $k$ ) is forbidden. Therefore, all these classes are defined by finitely many (not necessarily bipartite) forbidden induced subgraphs.

- $P_k$ -free bipartite permutation graphs. In [Korpelainen and Lozin, 2011b], it was shown that graphs in this class are  $k$ -letter graphs. Therefore, they are well-quasi-ordered by the labelled induced subgraph relation.
- $(P_7, Sun_1)$ -free bipartite graphs. Well-quasi-orderability of this class was also shown in [Korpelainen and Lozin, 2011b]. With no extra work this result can be extended to the labelled induced subgraph relation, because, as was shown in [Korpelainen and Lozin, 2011b], every connected graph in this class is  $C_4$ -free and  $(P_7, C_4)$ -free bipartite graphs contain no  $P_9$  as a (not necessarily

induced) subgraph. Now the conclusion follows immediately from the result of Ding [1992] discussed in Section 3.2.3. In Section 3.3, we will further extend this conclusion to more general classes. In particular, we will prove that the class of  $(P_k, \text{Sun}_2)$ -free bipartite graphs is well-quasi-ordered by the *labelled* induced subgraph relation for each fixed value of  $k$ . Notice that the class of  $(P_k, \text{Sun}_4)$ -free bipartite graphs contains an infinite antichain even for  $k = 7$  [Korpelainen and Lozin, 2011b], which leaves an interesting open question of determining well-quasi-orderability of  $(P_7, \text{Sun}_3)$ -free bipartite graphs.

- $(P_7, S_{1,2,3})$ -free bipartite graphs. This class was shown to be well-quasi-ordered by induced subgraphs in [Korpelainen and Lozin, 2011b]. It is known that  $(P_7, S_{1,2,3})$ -free bipartite graphs are totally decomposable by canonical decomposition [Fouquet et al., 1999] (meaning that the only indecomposable graph in this class is  $K_1$ ). Combining this fact with Theorem 3.1.3 we obtain the following conclusion.

**Corollary 3.2.2.** *The class of  $(P_7, S_{1,2,3})$ -free bipartite graphs is well-quasi-ordered by the labelled induced subgraph relation.*

### Bigenic classes

A hereditary class is *monogenic* if it is defined by one forbidden induced subgraph and *bigenic* if it is defined by two forbidden induced subgraphs. In the family of monogenic classes, the question of well-quasi-orderability by induced subgraphs was solved by Damaschke [1990] who showed that the class of graphs defined by a single forbidden induced subgraph  $H$  is well-quasi-ordered if and only if  $H$  is a (not necessarily proper) induced subgraph of  $P_4$ . It is known that every  $P_4$ -free graph with at least two vertices is either disconnected or the complement to a disconnected graph. Therefore, there are only two prime graph in this class:  $K_2$  and its complement. Therefore, according to Theorem 3.1.2, the class of  $P_4$ -free graphs is well-quasi-ordered by the labelled induced subgraph relation. Moreover, according to Damaschke [1990], this is the only maximal monogenic class satisfying this property.

In the family of bigenic classes, the situation is more complicated and it was studied in [Korpelainen and Lozin, 2011a]. This paper revealed a number of wqo and not-wqo bigenic classes and reduced the problem to 13 bigenic classes for which the question is open. We answer some of these questions in Section 3.3. Now we list the set of all maximal infinite bigenic classes for which well-quasi-orderability was shown earlier.



- $(K_3, P_3 + 2K_1)$ -free graphs. Bipartite graphs in this class are  $(P_7, S_{1,2,3})$ -free and hence are well-quasi-ordered by the *labelled* induced subgraph relation, as will be shown in Section 3.1.5. Also, in [Korpelainen and Lozin, 2011a] it was shown that non-bipartite graphs in this class are  $k$ -uniform for a finite value of  $k$ , i.e. they are well-quasi-ordered by *labelled* induced subgraphs too (Theorem 3.2.2).
- $(K_3, P_4 + K_1)$ -free graphs. Similarly to the previous case, bipartite graphs in this class are  $(P_7, S_{1,2,3})$ -free and non-bipartite are  $k$ -uniform for a finite  $k$  [Korpelainen and Lozin, 2011a], i.e. graphs in this class are well-quasi-ordered by the *labelled* induced subgraph relation.
- $(K_3, P_3 + P_2)$ -free graphs. As in the previous two cases, graphs in this class are well-quasi-ordered by the *labelled* induced subgraph relation, because they are either  $(P_7, S_{1,2,3})$ -free bipartite or  $k$ -uniform for a finite  $k$  [Korpelainen and Lozin, 2011a].
- $(P_5, \text{diamond})$ -free graphs. Every prime graph in this class is either  $k$ -uniform or a  $k$ -letter graph for a finite  $k$  [Korpelainen and Lozin, 2011a]. Together with Theorems 3.2.1, 3.2.2 and 3.1.2 this shows that Conjecture 3.1.1 is valid for this class.
- $(\text{diamond}, \text{co-diamond})$ -free graphs. This class is similar to the previous case, i.e. prime graphs in this class are either  $k$ -uniform or  $k$ -letter graphs for a finite  $k$  [Korpelainen and Lozin, 2011a]. Therefore, graphs in this class are well-quasi-ordered by the *labelled* induced subgraph relation, i.e. Conjecture 3.1.1 is valid for this class too.

## Two more examples

In this section, we mention two exotic examples of well-quasi-ordered classes of graphs. The first example deal with so called  *$\ell$ -dense graphs*. A graph with  $n$  vertices was called in [Zverovich, 2004]  $\ell$ -dense if it contains a clique with at least  $n - \ell$  vertices. The author proves that, for each fixed  $\ell \geq 0$ , the family of  $\ell$ -dense graphs is well-quasi-ordered with respect to the induced subgraph relation. It is not difficult to see that this is a particular example of  $k$ -uniform graphs, i.e. classes of  $\ell$ -dense graphs are finitely defined and well-quasi-ordered by labelled induced subgraphs.

The second example was introduced in [Ding, 1993]. It deals with graphs of bounded matroidal number. Let  $Y_{r,m}$  the class of graphs such that deletion of at most

$r$  vertices results in a graph of matroidal number at most  $m$ . It was shown in [Ding, 1993] that for each fixed  $r$ , the class  $Y_{r,3}$  is defined by finitely many forbidden induced subgraphs and is well-quasi-ordered by the labelled induced subgraph relation.

### 3.3 New results

In this section, we prove a number of new results that deal with classes of graphs defined by finitely many forbidden induced subgraphs. It is not difficult to see that such classes are well-quasi-ordered by induced subgraphs only if the set of forbidden graphs includes a path (or more generally, a linear forest), since otherwise they contain infinitely many cycles.

We start by presenting the following general result that comes readily from the following two facts:

- the class of graphs containing no  $P_k$  as a (not necessarily induced) subgraph is well-quasi-ordered by the *labelled* induced subgraph relation [Ding, 1992];
- every graph containing a large path as a (not necessarily induced) subgraph, contains either a large *induced* path or a large complete graph or a large complete bipartite graph as an induced subgraph proved in Chapter 6.

Together the two facts lead to the following conclusion.

**Theorem 3.3.1.** *For every fixed  $k, \ell$  and  $m$ , the class of  $(P_k, K_\ell, K_{m,m})$ -free graphs is well-quasi-ordered by the labelled induced subgraph relation.*

The remaining results in this section deal with more specific classes of graphs.

#### 3.3.1 Subclasses of $P_6$ -free graphs

In this section, we deal with bigenic subclasses of  $P_6$ -free graphs, i.e. subclasses defined by forbidding, in addition to  $P_6$ , one more induced subgraph. According to [Korpelainen and Lozin, 2011a], there are only two classes in this family for which the question of well-quasi-orderability by induced subgraphs is open. We answer this question for both them. In particular, we prove that the class of  $(P_6, K_3)$ -free graphs is well-quasi-ordered by the labelled induced subgraph relation, while the class of  $(P_6, \text{diamond})$ -free graphs contains an infinite antichain with respect to induced subgraphs.

### $(P_6, K_3)$ -free graphs

We start by quoting the following structural characterization of *prime*  $(P_6, K_3)$ -free graphs obtained in [Brandstädt et al., 2006].

**Theorem 3.3.2.** *Let  $G$  be a prime  $(P_6, K_3)$ -free graph which is not bipartite. Then the vertices of  $G$  can be partitioned into at most eleven sets  $X_0, X_1, \dots, X_5, Z_0, Z_1, \dots, Z_5$  and  $X_T$  such that*

- $|X_T| \leq 11$ ,
- $X_i \cup Z_i$  for  $i = 1, \dots, 5$  induce a bipartite graph, denoted  $B_i$ ,
- the graph  $G_0 = G[X_0 \cup Z_0]$  is either bipartite or has the following structure:  $G[Z_0]$  is a co-matched bipartite graph and  $X_0$  is an independent set whose vertices can be partitioned into two subsets  $X_{00}$  and  $X_{01}$  so that the vertices of  $X_{00}$  are isolated in  $G_0$  and each vertex of  $X_{01}$  has degree 2 in  $G_0$  and is adjacent to the endpoints of a non-edge in the co-matching of  $G[Z_0]$  (at most one vertex of  $X_{01}$  for each non-edge of  $G[Z_0]$ ).

The graphs  $G_0, B_1, \dots, B_5$  are called *basic*. Moreover, the sets  $X_0, X_1, \dots, X_5, Z_0, Z_1, \dots, Z_5$  can be further partitioned into at most 50 basic subsets so that any pair of basic subsets belonging to different basic graphs is either complete or anti-complete to each other.

Now we use this characterization of  $(P_6, K_3)$ -free graphs in order to prove the following theorem.

**Theorem 3.3.3.** *The class of  $(P_6, K_3)$ -free graphs is well-quasi-ordered by the labelled induced subgraph relation.*

*Proof.* From Corollary 3.2.2 we know that bipartite  $P_6$ -free graphs are well-quasi-ordered by the labelled induced subgraph relation (since  $P_6$  is an induced subgraph of both  $P_7$  and  $S_{1,2,3}$ ). To prove the result for non-bipartite graphs in our class we use Theorem 3.1.2, which allows us to restrict ourselves to prime graphs, and Theorem 3.3.2, which provides a characterization of prime non-bipartite graphs in the class.

Let  $G$  be a labelled prime non-bipartite  $(P_6, K_3)$ -free graph with vertex labels coming from a well-quasi-ordered set  $W$ . In addition to the original labels of  $G$  we attach to each vertex  $v$  of this graph a label  $l(v)$  defined as follows. The eleven vertices of  $X_T$  are labelled by numbers from 1 to 11. If  $v$  belongs to a basic subgraph, we define  $l(v) = (s_v, x_v)$ , where  $s_v$  is the number indicating the basic subset  $v$

belongs to (i.e. a number between 1 and 50) and  $x_v$  is a binary sequence of length 11 describing the neighbours of  $v$  in the set  $X_T$  (one bit for each vertex of  $X_T$ ). Since  $l(v)$  belongs to a finite set (of size at most  $50 \cdot 2^{11}$ ), the resulting label of each vertex of  $G$  (the original label plus  $l(v)$ ) belongs to a well-quasi-ordered set.

We describe the adjacency between different basic subsets of  $G$  with a binary sequence  $N$  of length  $\binom{50}{2}$  (one bit for each pair of subsets). Thus, the graph  $G$  can be completely described by the following 8-tuple  $(G[X_T], G_0, B_1, B_2, B_3, B_4, B_5, N)$ . Let us show that each entry of this tuple belongs to a well-quasi-ordered set. Each of the first seven entries is a labelled graph with labels coming from a well-quasi-ordered set. Since  $G[X_t]$  is finite and  $B_1, B_2, B_3, B_4, B_5$  are bipartite, each of them belongs to a well-quasi-ordered set. The graph  $G_0$  is either bipartite or 3-uniform (take  $F_k$  to be a triangle and  $K$  a  $3 \times 3$  matrix with 1's in positions  $(1, 2)$  and  $(2, 1)$  and 0's everywhere else), and hence belongs to a well quasi-ordered set by Theorem 3.2.2 and Corollary 3.2.2. The last entry belongs to a finite set, which is definitely well-quasi-ordered. Therefore, by Higman's lemma, the set of prime non-bipartite  $(P_6, K_3)$ -free graphs is well-quasi-ordered by the labelled induced subgraph. This completes the proof of the theorem.  $\square$

#### **$(P_6, \text{diamond})$ -free graphs are not WQO**

**Theorem 3.3.4.** *The class of  $(P_6, \text{diamond})$ -free graphs is not well-quasi-ordered by the induced subgraph relation.*

*Proof.* Let  $C_{4n}$  be a chordless cycle with  $4n$  vertices  $x_1, x_2, \dots, x_{4n}$ . We split the vertex set of this cycle into 3 subsets  $X, Y, Z$  as follows: let

- $X$  be the set of even-indexed vertices,
- $Y$  be the set of vertices  $x_i$  with index  $i \equiv 1 \pmod{4}$ ,
- $Z$  be the set of vertices  $x_i$  with index  $i \equiv 3 \pmod{4}$ .

Let us denote by  $G_{4n}$  the graph obtained from  $C_{4n}$  by connecting every vertex of  $Y$  to every vertex of  $Z$ . We will show that the set  $\{G_{4n} : n = 2, 3, 4, \dots\}$  is an infinite antichain with respect to the induced subgraph relation and every graph in this set is  $(P_6, \text{diamond})$ -free.

Firstly, observe that  $X, Y$  and  $Z$  are independent sets, so any triangle in  $G_{4n}$  must have exactly one vertex in each of them. As all vertices in  $X$  have degree 2, we conclude that the only triangles in  $G_{4n}$  are  $x_{2i-1}, x_{2i}, x_{2i+1}$  for  $i = 1, 2, \dots, 2n-1$  and  $x_{4n-1}, x_{4n}, x_1$ . Therefore, any two triangles in  $G_{4n}$  share at most one vertex,

and hence,  $G_{4n}$  is *diamond*-free for any  $n \geq 2$ . Moreover, the triangles of  $G_{4n}$  constitute a cyclic structure. More precisely, the edges of these triangles whose endpoints belong to  $Y$  and  $Z$  form a cycle. This cycle may contain chords but these chords belong to no triangles. Therefore, no  $G_{4n}$  can be embedded into a  $G_{4m}$  with  $n < m$  as an induced subgraph, because the cycle formed by the  $2n$  edges of the triangles of  $G_{4n}$  cannot be embedded into the cycle formed by the  $2m$  edges of the triangles of  $G_{4m}$ . This proves that  $\{G_{4n} : n = 2, 3, 4, \dots\}$  is an infinite antichain with respect to the induced subgraph relation.

Finally, take any induced path  $P$  of length 6 in  $G_{4n}$ . Then every vertex of  $P$  belonging to  $X$  must be an endpoint of  $P$ , since otherwise  $P$  contains a triangle. Therefore,  $P$  contains at most two vertices in  $X$  and hence at least four vertices in  $Y \cup Z$ . If two of these vertices belong to  $Y$  and two to  $Z$ , then  $P$  contains a  $C_4$ , which is impossible. If one vertex belongs to  $Y$  and three to  $Z$ , then  $P$  has a vertex of degree 3, which is impossible too. The remaining case when  $P$  has no vertices, say, in  $Z$  is impossible either, because the subgraph induced by  $X$  and  $Y$  is bipartite and in a bipartite graph any path of length 6 must contain 3 vertices in each part of the graph.  $\square$

### 3.3.2 $(P_k, Sun_2)$ -free bipartite graphs

A  $Sun_2$  is the graph represented in Figure 3.1. In this section, we prove that the class of  $(P_k, Sun_2)$ -free bipartite graphs is well-quasi-ordered by the *labelled* induced subgraph relation for each fixed value of  $k$ . Notice that the class of  $(P_k, Sun_4)$ -free bipartite graphs contains an infinite antichain even for  $k = 7$  [Korpelainen and Lozin, 2011b], which leaves an interesting open question of determining well-quasi-orderability of  $(P_7, Sun_3)$ -free bipartite graphs.

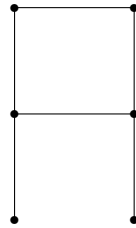


Figure 3.1: The graph  $Sun_2$

**Theorem 3.3.5.** *For each fixed value of  $k$ , the class of  $(P_k, Sun_2)$ -free bipartite graphs is well-quasi-ordered by the labelled induced subgraph relation.*

*Proof.* The set of  $(P_k, \text{Sun}_2)$ -free bipartite graphs containing no  $C_4$  as an induced subgraph is well-quasi-ordered under the labelled induced subgraph relation by Theorem 3.3.1. Therefore, we consider only those graphs in the class that contain a  $C_4$ . We will show that prime graphs of this type are 4-uniform.

Let  $G$  be a prime  $(P_k, \text{Sun}_2)$ -free bipartite graph containing a  $C_4$ . We extend this  $C_4$  to a maximal (with respect to set inclusion) complete bipartite graph  $H$  with parts  $A$  and  $B$ . Observe that  $|A| \geq 2$  and  $|B| \geq 2$ , since  $H$  contains a  $C_4$ . We denote by  $C$  the set of neighbours of  $B$  outside  $A$  (i.e. the set of vertices outside  $A$  each of which has at least one neighbour in  $B$ ) and by  $D$  the set of neighbours of  $A$  outside  $B$ . Notice that

- (1)  $C$  and  $D$  are non-empty, since otherwise  $A$  or  $B$  is a non-trivial module, contradicting the primality of  $G$ ;
- (2) each vertex of  $C$  has a non-neighbour in  $B$  and each vertex of  $D$  has a non-neighbour in  $A$  due to the maximality of  $H$ .

We also claim that

- (3)  $C \cup D$  induce a complete bipartite graph. Indeed, assume there are two non-adjacent vertices  $c \in C$  and  $d \in D$ . Consider a neighbour  $b_1$  and a non-neighbour  $b_2$  of  $c$  in  $B$ , and a neighbour  $a_1$  and a non-neighbour  $a_2$  of  $d$  in  $A$ . Then the six vertices  $a_1, a_2, b_1, b_2, c, d$  induce a  $\text{Sun}_2$  in  $G$ , a contradiction.
- (4)  $V(G) = A \cup B \cup C \cup D$ . To show this, assume there is a vertex  $x \notin A \cup B \cup C \cup D$ . With loss of generality we may assume that  $x$  is adjacent to a vertex  $c \in C$  (since  $G$  is prime and hence is connected). Let  $d$  be any vertex of  $D$ ,  $b$  any neighbour of  $c$  in  $B$ , and  $a_1, a_2$  a neighbour and a non-neighbour of  $d$  in  $A$ . Then the six vertices  $a_1, a_2, b, c, d, x$  induce a  $\text{Sun}_2$  in  $G$ , a contradiction.
- (5) every vertex of  $D$  has at most one non-neighbour in  $A$ . Assume, by contradiction, that a vertex  $d \in D$  has two non-neighbours  $a_1, a_2$  in  $A$ . Since  $G$  is prime, there must exist a vertex distinguishing  $a_1$  and  $a_2$  (otherwise  $\{a_1, a_2\}$  is a non-trivial module). Let  $d'$  be such a vertex. Clearly,  $d'$  belongs to  $D$ . Finally, consider any vertex  $c \in C$  and any of its neighbours  $b \in B$ . Then the six vertices  $a_1, a_2, b, c, d, d'$  induce a  $\text{Sun}_2$  in  $G$ , a contradiction.
- (6) every vertex of  $A$  has at most one non-neighbour in  $D$ . Assume, to the contrary, that a vertex  $a \in A$  has two non-neighbours  $d_1, d_2$  in  $D$ . Then, by (3) and (5),  $a$  is the only non-neighbour of  $d_1$  and  $d_2$ . But then  $\{d_1, d_2\}$  is a non-trivial module, contradicting the primality of  $G$ .

Claims (5) and (6) show that the bipartite complement of  $G[A \cup D]$  is a graph of vertex degree at most 1. Moreover, in this graph at most one vertex of  $A$  and at most one vertex of  $D$  have degree less than 1 (since  $G$  is prime). By symmetry, the bipartite complement of  $G[B \cup C]$  is a graph of degree at most 1 with at most one vertex of degree 0 in each part. But then  $G$  is a 4-uniform graph. Combining this fact with Theorems 3.2.2 and 3.1.2, we conclude that the set of  $(P_k, \text{Sun}_2)$ -free bipartite graphs containing a  $C_4$  is well-quasi-ordered by the labelled induced subgraph relation.  $\square$

### 3.4 Conclusion

In this chapter we worked with finitely defined hereditary classes of graphs comparing induced and labelled induced subgraph relations. We showed that all classes known to be well-quasi-ordered by induced subgraph relation are well-quasi-ordered by labelled induced relation as well. Whether these two notions coincide for all finitely defined hereditary classes of graphs remains an open question.

In addition, in Section 3.3 we revealed some new classes supporting the conjecture. In particular, we proved that the graphs forbidding  $P_6$  and  $K_3$  are well-quasi-ordered by the induced subgraph relation. This result is interesting because it completes the question of deciding well-quasi-ordering for bigenic graph classes forbidding a clique and a path. Indeed, on the side of wqo, graphs avoiding  $K_2$  are trivially wqo and graphs avoiding  $P_4$  (co-graphs) are well-known to be wqo and our result proves that classes forbidding  $K_3$  and  $P_5$  or  $P_6$  are well-quasi-ordered. On the side of non-wqo, two infinite antichains have already been constructed: one (from [Korpelainen and Lozin, 2011a]) in the class of graphs with neither  $P_5$  (or even  $2K_2$ ) nor  $K_4$ , and one (from [Korpelainen and Lozin, 2011b]) in the class omitting both  $P_7$  and  $K_3$ . The construction of antichains for these classes used a certain correspondence between antichains for permutations and graphs. In the next section we will consider classes of permutation graphs excluding a path and a clique and we will reveal a new family of finitely defined classes which are well-quasi-ordered by the induced subgraph relation. Whether these classes are well-quasi-ordered by the labelled induced subgraph relation is the next interesting research question.

## Chapter 4

# Permutations and permutation graphs

In this chapter we consider finite graphs and permutation graphs which omit both a path  $P_k$  and a clique  $K_\ell$ . Our main result, proved in Section 4.2, is that permutation graphs which avoid both  $P_5$  and a clique  $K_\ell$  are well-quasi-ordered under the induced subgraph order for every finite  $\ell$ . We also prove, in Section 4.3, that the three classes of permutation graphs defined by forbidding  $\{P_6, K_6\}$ ,  $\{P_7, K_5\}$  and  $\{P_8, K_4\}$  respectively are not well-quasi-ordered, by exhibiting an infinite antichain in each case.

A summary of our results is shown in Figure 4.1, with our new contributions in the upper-right highlighted. Filled circles indicate that all graphs avoiding the specified path and clique are wqo. Half-filled circles indicate that the corresponding class of permutation graphs are wqo, but that the corresponding class of all graphs are not wqo. Empty circles indicate that neither class is wqo. Note that for the three unknown cases (indicated by question marks), it is known that the corresponding class of graphs contains an infinite antichain.

We start by setting up the necessary notions from the study of permutation classes.

### 4.1 Definitions and structural tools

Given a permutation  $\pi = \pi(1) \cdots \pi(n)$ , its corresponding *permutation graph* is the graph  $G_\pi$  on the vertices  $\{1, \dots, n\}$  in which  $i$  is adjacent to  $j$  if both  $i \leq j$  and  $\pi(i) \geq \pi(j)$ . This mapping is many-to-one, because, for example,  $G_{231} \cong G_{312} \cong P_3$ . Given permutations  $\sigma = \sigma(1) \cdots \sigma(k)$  and  $\pi = \pi(1) \cdots \pi(n)$ , we say that  $\sigma$  is



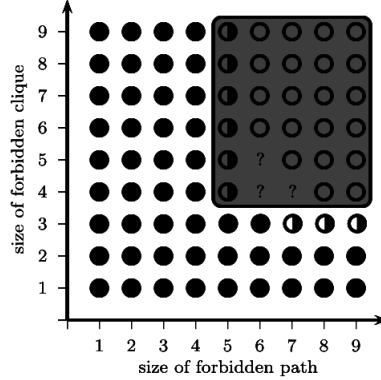


Figure 4.1: WQO results for classes of graphs and permutation graphs avoiding paths and cliques

*contained* in  $\pi$  and write  $\sigma \leq \pi$  if there are indices  $1 \leq i_1 < \dots < i_k \leq n$  such that the sequence  $\pi(i_1) \dots \pi(i_k)$  is in the same relative order as  $\sigma$ . For us, a *class* of permutations is a set of permutations closed downward under this containment order. Each such class can be represented by the minimal set of permutations not contained in the class, and we denote the class of permutations "avoiding" (i.e. not containing) the given set  $S$  of permutations by  $\text{Av}(S)$ . We say that a class or permutations is wqo if it does not contain an infinite antichain.

The mapping  $\pi \mapsto G_\pi$  is easily seen to be *order-preserving*, i.e., if  $\sigma \leq \pi$  then  $G_\sigma$  is an induced subgraph of  $G_\pi$ . Therefore if a class  $\mathcal{C}$  of permutations is wqo then the associated class of permutation graphs  $\{G_\pi : \pi \in \mathcal{C}\}$  must also be wqo. However, there are infinite antichains of permutations which do not correspond to antichains of permutation graphs. For this reason, when showing that classes of permutation graphs are wqo we instead prove the stronger result that the associated permutation classes are wqo, but when constructing infinite antichains, we must construct antichains of permutation graphs.

Instead of working directly with permutation graphs, we establish our wqo results for the corresponding permutation classes (which, by our observations in the introduction, is a stronger result). The permutations 24153 and 31524 are the only permutations which correspond to the permutation graph  $P_5$ , and thus to establish our main result we must determine the structure of the permutation class

$$\text{Av}(24153, 31524, \ell \dots 21).$$

Considering the wqo problem from the viewpoint of permutations has the added benefit of allowing us to make use of the recently developed tools in this field. In particular, we utilise grid classes and the substitution decomposition. Thus before establishing our main result we must first introduce these concepts.

We frequently identify a permutation  $\pi$  of length  $n$  with its *plot*, the set  $\{(i, \pi(i)) : 1 \leq i \leq n\}$  of points in the plane. We say that a rectangle in the plane is *axis-parallel* if its top and bottom sides are parallel with the  $x$ -axis while its left and right sides are parallel with the  $y$ -axis. Given natural numbers  $i$  and  $j$  we denote by  $[i, j]$  the closed interval  $\{i, i+1, \dots, j\}$  and by  $[i, j)$  the half-open interval  $\{i, i+1, \dots, j-1\}$ . Thus the axis-parallel rectangles we are interested in may be described by  $[x, x'] \times [y, y']$  for natural numbers  $x, x', y$ , and  $y'$ .

Monotone grid classes are a way of partitioning the entries of a permutation (or rather, its plot) into monotone axis-parallel rectangles in a manner specified by a  $0/\pm 1$  matrix. In order for these matrices to align with plots of permutations, we index them with Cartesian coordinates. Suppose that  $M$  is a  $t \times u$  matrix (thus  $M$  has  $t$  columns and  $u$  rows). An  $M$ -*gridding* of the permutation  $\pi$  of length  $n$  consists of a pair of sequences  $1 = c_1 \leq \dots \leq c_{t+1} = n+1$  and  $1 = r_1 \leq \dots \leq r_{u+1} = n+1$  such that for all  $k$  and  $\ell$ , the entries of (the plot of)  $\pi$  that lie in the axis-parallel rectangle  $[c_k, c_{k+1}) \times [r_\ell, r_{\ell+1})$  are increasing if  $M_{k,\ell} = 1$ , decreasing if  $M_{k,\ell} = -1$ , or empty if  $M_{k,\ell} = 0$ .

We say that the permutation  $\pi$  is  $M$ -*griddable* if it possesses an  $M$ -gridding, and the *grid class* of  $M$ , denoted by  $\text{Grid}(M)$ , consists of the set of  $M$ -griddable permutations. We further say that the permutation class  $\mathcal{C}$  is  $M$ -griddable if  $\mathcal{C} \subseteq \text{Grid}(M)$ , and that this class is *monotone griddable* if there is a finite matrix  $M$  for which it is  $M$ -griddable.

Grid classes were first described in this generality (albeit under a different name) by Murphy and Vatter [2003], who studied their wqo properties. To describe their result we need the notion of the *cell graph* of a matrix  $M$ . This graph has vertex set  $\{(i, j) : M_{i,j} \neq 0\}$  and  $(i, j)$  is adjacent to  $(k, \ell)$  if they lie in the same row or column and there are no nonzero entries lying between them in this row or column. We typically attribute properties of the cell graph of  $M$  to  $M$  itself; thus we say that  $M$  is a forest if its cell graph is a forest.

**Theorem 4.1.1** ([Murphy and Vatter, 2003]). *The grid class  $\text{Grid}(M)$  is wqo if and only if  $M$  is a forest.*

(This is a slightly different form of the result than is stated in [Murphy and Vatter, 2003], but the two forms are equivalent as shown by Vatter and Waton [2011], who also gave a much simpler proof of Theorem 4.1.1.)

The monotone griddable classes were characterised by Huczynska and Vatter [2006]. In order to present this result we also need some notation. Given permutations  $\sigma$  and  $\tau$  of respective lengths  $m$  and  $n$ , their *sum* is the permutation  $\pi \oplus \tau$  whose plot consists of the plot of  $\tau$  above and to the right of the plot of  $\sigma$ . More formally, this permutation is defined by

$$\sigma \oplus \tau(i) = \begin{cases} \sigma(i) & \text{for } 1 \leq i \leq m, \\ \tau(i - m) + m & \text{for } m + 1 \leq i \leq m + n. \end{cases}$$

The obvious symmetry of this operation (in which the plot of  $\tau$  lies below and to the right of the plot of  $\sigma$ ) is called the *skew sum* of  $\sigma$  and  $\tau$  and is denoted  $\sigma \ominus \tau$ . We can now state the characterisation of monotone griddable permutation classes.

**Theorem 4.1.2** ([Huczynska and Vatter, 2006]). *The permutation  $\mathcal{C}$  is monotone griddable if and only if it does not contain arbitrarily long sums of 21 or skew sums of 12.*

The reader might note that the classes we are interested in are not monotone griddable, let alone  $M$ -griddable for a forest  $M$ . However, our proof will show that these classes can be built from a monotone griddable class via the “substitution decomposition”, which we define shortly. Before this, though, we must introduce a few more concepts concerning monotone grid classes, the first two of which are alternative characterisations of monotone griddable classes.

Given a permutation  $\pi$ , we say that the axis-parallel rectangle  $R$  is *monotone* if the entries of  $\pi$  which lie in  $R$  are monotone increasing or decreasing (otherwise  $R$  is *non-monotone*). We say that the permutation  $\pi$  can be *covered by  $s$  monotone rectangles* if there is a collection  $\mathfrak{R}$  of  $s$  monotone axis-parallel rectangles such that every point in the plot of  $\pi$  lies in at least one rectangle in  $\mathfrak{R}$ . Clearly if  $\mathcal{C} \subseteq \text{Grid}(M)$  for a  $t \times u$  matrix  $M$  then every permutation in  $\mathcal{C}$  can be covered by  $tu$  monotone rectangles. To see the other direction, note that every permutation which can be covered by  $s$  monotone rectangles is  $M$ -griddable for some matrix  $M$  of size at most  $(2s - 1) \times (2s - 1)$ . There are only finitely many such matrices, say  $M^{(1)}, \dots, M^{(m)}$ , so their *direct sum*,

$$\begin{pmatrix} & & & M^{(m)} \\ & & & \\ & & \ddots & \\ & & & \\ M^{(1)} & & & \end{pmatrix}$$

is a finite matrix whose grid class contains all such permutations.<sup>1</sup>

---

<sup>1</sup>By adapting this argument it follows that if every permutation in the class  $\mathcal{C}$  lies in the grid

This characterisation of monotone griddability is recorded in Proposition 4.1.1 below, which also includes a third characterisation. We say that the line  $L$  *slices* the rectangle  $R$  if  $L \cap R \neq \emptyset$ . If  $\mathcal{C} \subseteq \text{Grid}(M)$  for a  $t \times u$  matrix then for every permutation  $\pi \in \mathcal{C}$  there is a collection of  $t + u$  horizontal and vertical lines (the grid lines) which slice every non-monotone axis-parallel rectangle of  $\pi$ . Conversely, every such collection of lines defines a gridding of  $\pi$ , completing the sketch of the proof of the following result.

**Proposition 4.1.1** (specialising Proposition 2.3 of [Vatter, 2011]). *For a permutation class  $\mathcal{C}$ , the following are equivalent:*

- (1)  $\mathcal{C}$  is monotone griddable,
- (2) there is a constant  $\ell$  such that for every permutation  $\pi \in \mathcal{C}$  the set of non-monotone axis-parallel rectangles of  $\pi$  can be sliced by a collection of  $\ell$  horizontal and vertical lines, and
- (3) there is a constant  $s$  such that every permutation in  $\mathcal{C}$  can be covered by  $s$  monotone rectangles.

We now move on to the substitution decomposition, which will allow us to build the classes we are interested in from grid classes of forests. An *interval* in the permutation  $\pi$  is a set of contiguous indices  $I = [a, b]$  such that the set of values  $\pi(I) = \{\pi(i) : i \in I\}$  is also contiguous. Given a permutation  $\sigma$  of length  $m$  and nonempty permutations  $\alpha_1, \dots, \alpha_m$ , the *inflation* of  $\sigma$  by  $\alpha_1, \dots, \alpha_m$  — denoted  $\sigma[\alpha_1, \dots, \alpha_m]$  — is the permutation of length  $|\alpha_1| + \dots + |\alpha_m|$  obtained by replacing each entry  $\sigma(i)$  by an interval that is order isomorphic to  $\alpha_i$  in such a way that the intervals themselves are order isomorphic to  $\sigma$ . Thus the sum and skew sum operations are particular cases of inflations:  $\sigma \oplus \tau = 12[\sigma, \tau]$  and  $\sigma \ominus \tau = 21[\sigma, \tau]$ . Given two classes  $\mathcal{C}$  and  $\mathcal{U}$ , the *inflation* of  $\mathcal{C}$  by  $\mathcal{U}$  is defined as

$$\mathcal{C}[\mathcal{U}] = \{\sigma[\alpha_1, \dots, \alpha_m] : \sigma \in \mathcal{C}_m \text{ and } \alpha_1, \dots, \alpha_m \in \mathcal{U}\}.$$

The class  $\mathcal{C}$  is said to be *substitution closed* if  $\mathcal{C}[\mathcal{C}] = \mathcal{C}$ . The *substitution closure*,  $\langle \mathcal{C} \rangle$ , of a class  $\mathcal{C}$  is defined as the smallest substitution closed class containing  $\mathcal{C}$ . A standard argument shows that  $\langle \mathcal{C} \rangle$  exists, and by specialising a result of Albert et al. [2014] we obtain the following.

**Theorem 4.1.3** (specialisation of Theorem 4.4 of [Albert et al., 2014]). *If the matrix  $M$  is a forest then the class  $\langle \text{Grid}(M) \rangle$  is wqo.*

---

class of a forest of size at most  $t \times u$ , then  $\mathcal{C}$  itself lies in the grid class of a (possibly much larger) forest.

The permutation  $\pi$  is said to be *simple* if it cannot be written as a nontrivial inflation. Thus if  $\pi$  has length  $n$ , it is simple if and only if its only intervals have length 0, 1, and  $n$ . Thus every permutation can be expressed as the inflation of a simple permutation. Moreover, in most cases, this decomposition is unique. A permutation is said to be *sum* (resp., *skew*) *decomposable* if it can be expressed as the sum (resp., skew sum) of two shorter permutations. Otherwise it is said to be *sum* (resp., *skew*) *indecomposable*.

**Proposition 4.1.2** ([Albert and Atkinson, 2005]). *Every permutation  $\pi$  except 1 is the inflation of a unique simple permutation  $\sigma$ . Moreover, if  $\pi = \sigma[\alpha_1, \dots, \alpha_m]$  for a simple permutation  $\sigma$  of length  $m \geq 4$ , then each interval  $\alpha_i$  is unique. If  $\pi$  is an inflation of 12 (i.e., is sum decomposable), then there is a unique sum indecomposable  $\alpha_1$  such that  $\pi = \alpha_1 \oplus \alpha_2$ . The same holds, mutatis mutandis, with 12 replaced by 21 and sum replaced by skew.*

We close this section by noting how easily this machinery can show that permutation graphs omitting both  $P_k$  and  $K_3$  are wqo for all  $k$ , a result originally due to Korpelainen and Lozin [2011b]. The corresponding permutations all avoid 321 and thus lie in the grid class of an *infinite* matrix, known as the infinite staircase (see [Albert and Vatter, 2013]). Moreover, the sum indecomposable permutations which avoid 321 and the two permutations corresponding to  $P_k$  can be shown to lie in the grid class of a finite staircase,

$$\begin{pmatrix} & & & 1 & 1 \\ & & \ddots & \ddots & \\ & 1 & 1 & & \\ 1 & 1 & & & \end{pmatrix}.$$

Finally, it follows by an easy application of Higman's Theorem [Higman, 1952] that if the sum indecomposable permutations in a class are wqo, then the class itself is wqo. (In this case, the wqo conclusion also follows by Theorem 4.1.3.)

## 4.2 Permutation Graphs Omitting $P_5$ and $K_\ell$

In this section we prove that the class of permutations corresponding to permutation graphs omitting  $P_5$  and  $K_\ell$ ,

$$\text{Av}(24153, 31524, \ell \cdots 21),$$

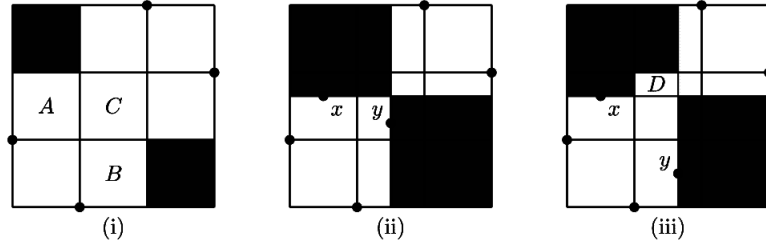


Figure 4.2: The impossible configuration for a simple permutation in Proposition 4.2.1.

is wqo. Our proof basically consists of two steps. First, we show that the simple permutations in these classes are monotone griddable, and then we show that these griddings can be refined to forests. The conclusion then follows from Theorem 4.1.3.

Given a set of points in the plane, their *rectangular hull* is defined to be the smallest axis-parallel rectangle containing all of them. We begin with a very simple observation about these simple permutations.

**Proposition 4.2.1.** *For every simple permutation  $\pi \in \text{Av}(24153, 31524)$ , either its greatest entry lies to the left of its least entry, or its leftmost entry lies above its rightmost entry.*

*Proof.* Suppose, for a contradiction, that  $\pi$  is a simple permutation in  $\text{Av}(24153, 31524)$  such that its greatest entry lies to the right of its least, and its leftmost entry lies below its rightmost entry. Thus, these four extremal entries form the pattern 2143, and the situation is as depicted in Figure 4.2(i). Since  $\pi$  is simple, regions  $A$  and  $B$  cannot both be empty, so, without loss of generality, suppose that  $A$  is non-empty and label the greatest entry in this region as point  $x$ .

Since  $\pi$  is simple, the rectangular hull of the leftmost entry, the least entry, and the point  $x$  cannot be an interval in  $\pi$ . Therefore, there must be a point either in  $B$ , or in that part of  $C$  lying below  $x$ . Take the rightmost such point, and label it  $y$ . If  $y$  is in region  $C$ , we immediately encounter the contradiction illustrated in Figure 4.2(ii): our choices of  $x$  and  $y$ , and the forbidden permutation 24153 causes the permutation to be sum decomposable. Therefore,  $y$  is placed in region  $B$ , and we have the picture depicted in Figure 4.2(iii). Since  $\pi$  is simple, there must now be a point in the region labelled  $D$ . However, labelling the greatest entry in this region yields another contradiction, as  $\pi$  must again be sum decomposable.  $\square$

The class  $\text{Av}(24153, 31524, \ell \cdots 21)$  is closed under group-theoretic inversion (because  $24153^{-1} = 31524$  and  $\ell \cdots 21^{-1} = \ell \cdots 21$ ), so we may always assume that

the latter option in Proposition 4.2.1 holds.

The rest of our proof adapts several ideas from Vatter [2011]. Two rectangles in the plane are said to be *dependent* if their projections onto either the  $x$ - or  $y$ -axis have nontrivial intersection, and otherwise they are said to be *independent*. A set of rectangles is called independent if its members are pairwise independent. Thus an independent set of rectangles may be viewed as a permutation, and it satisfies the Erdős-Szekeres Theorem (every permutation of length at least  $(a-1)(b-1)+1$  contains either  $12\cdots a$  or  $b\cdots 21$ ). We construct independent sets of rectangles in the proofs of both Propositions 4.2.2 and 4.2.4. In these settings, the rectangles are used to capture “bad” areas in the plot of a permutation, and our desired result is obtained by slicing the rectangles with horizontal and vertical lines in the sense of Proposition 4.1.1. The following result shows that we may slice a collection of rectangles with only a few lines, so long as we can bound its independence number.

**Theorem 4.2.1** ([Gyárfás and Lehel, 1970]). *There is a function  $f(m)$  such that for any collection  $\mathfrak{R}$  of axis-parallel rectangles in the plane which has no independent set of size greater than  $m$ , there exists a set of  $f(m)$  horizontal and vertical lines which slice every rectangle in  $\mathfrak{R}$ .*

Our next two results both rest upon Theorem 4.2.1.

**Proposition 4.2.2.** *For every  $\ell$ , the simple permutations in  $\text{Av}(24153, 31524, \ell \cdots 21)$  are contained in  $\text{Grid}(M)$  for a finite 0/1 matrix  $M$ .*

*Proof.* By Proposition 4.1.1 it suffices to show that there is a function  $g(\ell)$  such that every permutation in  $\text{Av}(24153, 31524, \ell \cdots 21)$  can be covered by  $g(\ell)$  increasing rectangles (i.e., rectangles which only cover increasing sets of points). We prove this statement by induction on  $\ell$ . For the base case, we can take  $f(2) = 1$ . Now take a simple permutation  $\pi \in \text{Av}(24153, 31524, \ell \cdots 21)$  for  $\ell \geq 3$  and suppose that the claim holds for  $\ell - 1$ .

By Proposition 4.2.1, we may assume that the leftmost entry of  $\pi$  lies above its rightmost entry. Let  $\pi_t$  be the permutation formed by all entries of  $\pi$  lying above its rightmost entry and  $\pi_b$  the permutation formed by all entries of  $\pi$  lying below its leftmost entry (as shown in Figure 4.3). Thus every entry of  $\pi$  corresponds to an entry in  $\pi_t$ , to an entry in  $\pi_b$ , or to entries in both permutations. Moreover, both  $\pi_t$  and  $\pi_b$  avoid  $(\ell - 1) \cdots 21$ .

We would like to apply induction to find monotone rectangle coverings of both  $\pi_t$  and  $\pi_b$  but of course these permutations needn’t be simple. Nevertheless, if  $\pi_t$  is the inflation of the simple permutation  $\sigma_t$  and  $\pi_b$  of  $\sigma_b$  then both  $\sigma_t$  and

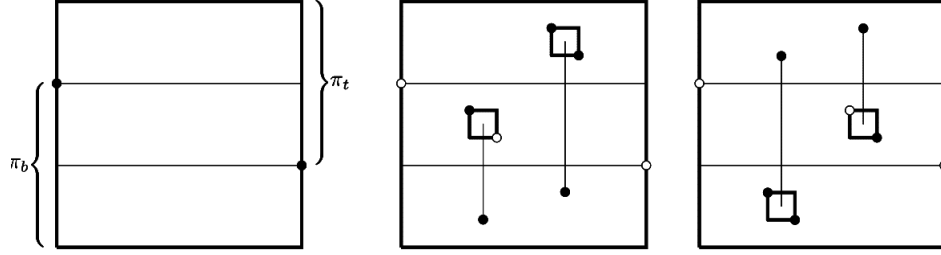


Figure 4.3: The division of  $\pi$  used in Proposition 4.2.2 (on the left) and the final contradictions in the proof (in the centre and on the right)

$\sigma_b$  can be covered by  $g(\ell - 1)$  increasing rectangles by induction. Now we stretch these increasing rectangles so that they cover the corresponding regions of  $\pi$ . By adapting the proof of Proposition 4.1.1, we may then extend this rectangle covering to a gridding of size at most  $4g(\ell - 1) \times 4g(\ell - 1)$ . While this gridding needn't be monotone, inside each of its cells we see points which correspond either to inflations of increasing sequences of  $\sigma_t$  or of  $\sigma_b$ . Let  $\mathfrak{L}$  denote the grid lines of this gridding.

We now say that the axis-parallel rectangle  $R$  is *bad* if it is fully contained in a cell of the above gridding and the points it covers contain a decreasing interval in either  $\pi_t$  or  $\pi_b$ . Further let  $\mathfrak{R}$  denote the collection of all bad rectangles. We aim to show that there is a collection of  $f(2\ell(\ell - 1))$  lines which slice every bad rectangle, where  $f$  is the function defined in Theorem 4.2.1. This, together with Proposition 4.1.1 and the comments before it, will complete the proof because these lines together with  $\mathfrak{L}$  will give a gridding of  $\pi$  of bounded size.

Theorem 4.2.1 will give us the desired lines if we can show that  $\mathfrak{R}$  has no independent set of size greater than  $2\ell(\ell - 1) + 1$ . Suppose to the contrary that  $\mathfrak{R}$  does contain an independent set of this size. Thus at least  $\ell(\ell - 1) + 1$  of these bad rectangles lie in one of  $\pi_t$  or  $\pi_b$ ; suppose first that these  $\ell(\ell - 1) + 1$  bad rectangles lie in  $\pi_t$ . Because  $\pi_t$  avoids  $(\ell - 1) \cdots 21$ , the Erdős-Szekeres Theorem implies that at least  $\ell + 1$  of its bad rectangles occur in increasing order (when read from left to right). Because  $\pi$  itself is simple, each such rectangle must be separated, and this separating point must lie in  $\pi_t \setminus \pi_b$ . Appealing once more to Erdős-Szekeres we see that two such separating points must themselves lie in increasing order, as shown in the centre of Figure 4.3. However, this is a contradiction to our assumption that  $\pi$  avoids 31524 (given by the solid points). As shown on the rightmost pane of this figure, if the  $\ell(\ell - 1) + 1$  bad rectangles lie in  $\pi_b$  we instead find a copy of 24153.  $\square$

A *submatrix* of a matrix is obtained by deleting any collection of rows and



columns from the matrix. Our next result shows that the simple permutations in  $\text{Av}(24153, 31524, \ell \cdots 21)$  can be gridded in a matrix which does not contain  $\begin{pmatrix} 1 & 1 \\ 1 & * \end{pmatrix}$  as a submatrix, i.e. in this matrix, there is no non-zero cell with both a non-zero cell below it in the same column, and a non-zero cell to its right in the same row. (The  $*$  indicates an entry that can be either 0 or 1.)

The following result is in some sense the technical underpinning of our entire argument. We advise the reader to note during the proof that if the hulls in  $\mathfrak{H}$  are assumed to be increasing, then the resulting gridding matrix  $M$  will be 0/1, not 0/ $\pm 1$ .

**Proposition 4.2.3.** *Suppose that  $\pi$  is a permutation and  $\mathfrak{H}$  is a collection of  $m$  monotone rectangles which cover the entries of  $\pi$  satisfying*

- (H1) *the hulls in  $\mathfrak{H}$  are pairwise nonintersecting,*
- (H2) *no single vertical or horizontal line slices through more than  $k$  hulls in  $\mathfrak{H}$ , and*
- (H3) *no hull in  $\mathfrak{H}$  is dependent both with a hull to its right and a hull beneath it.*

*Then there exists a function  $f(m, k)$  such that  $\pi$  is  $M$ -griddable for a 0/ $\pm 1$  matrix  $M$  of size at most  $f(m, k) \times f(m, k)$  which does not contain  $\begin{pmatrix} \pm 1 & \pm 1 \\ \pm 1 & * \end{pmatrix}$  as a submatrix.*

*Proof.* We define a gridding of  $\pi$  using the sides of the hulls in  $\mathfrak{H}$ . For a given side from a given hull, form a gridline by extending it to the edges of the permutation. By our hypothesis (H2), this line can slice at most  $k$  other hulls.

Whenever a hull is sliced in this way, a second gridline, perpendicular to the first, is induced so that all of the entries within the hull are contained in the bottom left and top right quadrants (for a hull containing increasing entries), or top left and bottom right quadrants (for a hull containing decreasing entries) defined by the two lines. This second gridline may itself slice through the interior of at most  $k$  further hulls in  $\mathfrak{H}$ , and each such slice will induce another gridline, and so on. We call this process the *propagation* of a line. See Figure 4.4 for an illustration. For a given propagation sequence, the *propagation tree* has gridlines for vertices, is rooted at the original gridline for the side, and has an edge between two gridlines if one induces the other in this propagation.

Before moving on, we note that it is clear that the propagation tree is connected, but less obvious that it is in fact a tree. This is not strictly required in our argument (we will only need to bound the number of vertices it contains), but if the tree were to contain a cycle it would have to correspond to a cyclic sequence of hulls, and this is impossible without contradicting hypothesis (H3).

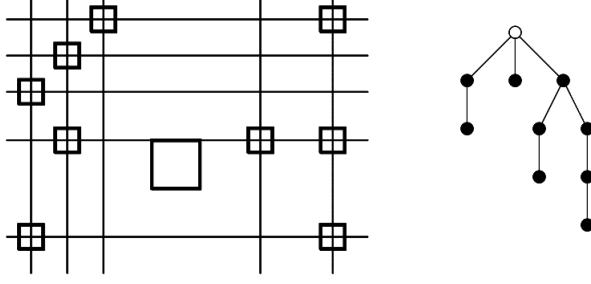


Figure 4.4: Propagating a side of a hull in  $\mathfrak{H}$ , and its corresponding propagation tree.

In order to bound the size of a propagation tree, we first show that it has height at most  $2m - 1$ . In the propagation tree of a side from some hull  $H_0 \in \mathfrak{H}$ , take a sequence  $H_1, H_2, \dots, H_p$  of hulls from  $\mathfrak{H}$  corresponding to a longest path in the propagation tree, starting from the root. Thus,  $H_1$  is sliced by the initial gridline formed from the side of  $H_0$ , and  $H_i$  is sliced by the gridline induced from  $H_{i-1}$  for  $i = 2, \dots, p$ .

We now define a word  $w = w_1 w_2 \dots w_p$  from this sequence, based on the position of hull  $H_i$  relative to hull  $H_{i-1}$ . For  $i = 1, 2, \dots, p$ , let

$$w_i = \begin{cases} \text{u} & \text{if } H_i \text{ lies above } H_{i-1} \\ \text{d} & \text{if } H_i \text{ lies below } H_{i-1} \\ \text{l} & \text{if } H_i \text{ lies to the left of } H_{i-1} \\ \text{r} & \text{if } H_i \text{ lies to the right of } H_{i-1} \end{cases}$$

Note that by the process of inducing perpendicular gridlines, successive letters in  $w$  must alternate between  $\{\text{u}, \text{d}\}$  and  $\{\text{l}, \text{r}\}$ . Moreover, since no hull interacts with hulls both below it and to its right,  $w$  cannot contain a  $\text{ur}$  or  $\text{ld}$  factor. In other words, after the first instance of  $\text{u}$  or  $\text{l}$ , there are no more instances of  $\text{r}$  or  $\text{d}$ . This means that  $w$  consists of a (possibly empty) alternating sequence  $\text{drdr} \dots$  or  $\text{rdrd} \dots$  followed by a (possibly empty) alternating sequence  $\text{ulul} \dots$  or  $\text{lulu} \dots$ .

Any alternating sequence of the form  $\text{drdr} \dots$  or  $\text{rdrd} \dots$  can have at most  $m - 1$  letters, as each hull in  $\mathfrak{H}$  (other than  $H_0$ ) can be sliced at most once in such a sequence. Similarly, any alternating sequence of the form  $\text{ulul} \dots$  or  $\text{lulu} \dots$  can have at most  $m - 1$  letters. Consequently, we have  $p \leq 2m - 2$ , and thus every propagation sequence has length at most  $2m - 1$ , as required.

Since each gridline in the propagation tree has at most  $k$  children, this means

that the propagation tree for any given side has at most

$$1 + k + k^2 + \dots + k^{2m-1} < k^{2m}$$

vertices, yielding a gridding of  $\pi$  with fewer than  $4mk^{2m}$  gridlines, and we may take this number to be  $f(m, k)$ . The gridding matrix  $M$  is then naturally formed from the cells of this gridding of  $\pi$ : each empty cell corresponds to a 0 in  $M$ , each cell containing points in decreasing order corresponds to a  $-1$  in  $M$ , and each cell containing points in increasing order corresponds to a 1 in  $M$ .

Finally, we verify that  $M$  satisfies the conditions in the proposition. The process of propagating gridlines ensures that each rectangular hull in  $\mathfrak{H}$  is divided into cells no two of which occupy the same row or column of the  $M$ -gridding. This means that there can be no  $\begin{pmatrix} \pm 1 & \pm 1 \\ \pm 1 & * \end{pmatrix}$  submatrix of  $M$  with two cells originating from the same hull in  $\mathfrak{H}$ . Thus, the cells of a submatrix of the form  $\begin{pmatrix} \pm 1 & \pm 1 \\ \pm 1 & * \end{pmatrix}$  must be made up from points in distinct hulls in  $\mathfrak{H}$ , but this is impossible since  $\mathfrak{H}$  contains no hulls which are dependent with hulls both below it and to its right.  $\square$

We now apply this proposition to refine the gridding provided to us by Proposition 4.2.2.

**Proposition 4.2.4.** *For every  $\ell$ , the simple permutations in  $\text{Av}(24153, 31524, \ell \dots 21)$  are contained in  $\text{Grid}(M)$  for a finite 0/1 matrix  $M$  which does not contain  $\begin{pmatrix} 1 & 1 \\ 1 & * \end{pmatrix}$  as a submatrix.*

*Proof.* Let  $\pi$  be an arbitrary simple permutation in  $\text{Av}(24153, 31524, \ell \dots 21)$ . As we observed in Footnote 1, it suffices to show that there are constants  $a$  and  $b$  such that  $\pi \in \text{Grid}(M)$  for an  $a \times b$  matrix  $M$  satisfying the desired conditions.

By Proposition 4.2.2,  $\pi$  is contained in  $\text{Grid}(N)$  for some 0/1 matrix  $N$ , of size (say)  $t \times u$ . We say that a *bad* rectangle within any specified cell is an axis-parallel rectangle which contains two entries which are split both by points below and to the right. Since  $\pi$  does not contain 24153, no cell can contain more than one independent bad rectangle — see Figure 4.5 (the filled points in the Figure form a copy of 24153, irrespective of the relative orders of the splitting points). Therefore an independent set of bad rectangles can have size at most  $tu$ , so Theorem 4.2.1 shows that the bad rectangles can all be sliced by  $f(tu)$  lines.

In any cell of the original gridding, the additional  $f(tu)$  slices that have been added can at most slice these points into  $f(tu) + 1$  maximal unsliced pieces. In the entire permutation, therefore, these slices (together with the original gridlines from

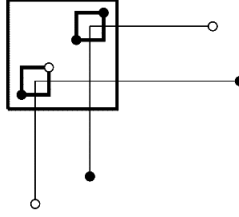


Figure 4.5: Two independent bad rectangles

$N$ ) divide the points into at most  $tu(f(tu) + 1)$  maximal unsliced pieces. Let  $\mathfrak{H}$  denote the rectangular hulls of these maximal unsliced pieces.

We now check that  $\mathfrak{H}$  satisfies the hypotheses of Proposition 4.2.3. Condition (H1) follows immediately by construction. Next, note that no two hulls of  $\mathfrak{H}$  from the same cell of the  $N$ -gridding can be dependent, since every cell is monotone, so we may take  $k = \max(t, u)$  to satisfy (H2). Finally, no hull can contain a bad rectangle (since all bad rectangles have been sliced), and so no hull can simultaneously be dependent with a hull from a cell below it, and a hull from a cell to its right, as required by (H3).

Now, applying Proposition 4.2.3 (noting that all the rectangular hulls in  $\mathfrak{H}$  contain increasing entries), we have a 0/1 gridding matrix  $M_\pi$  for  $\pi$ , of dimensions at most  $v \times w$  for some  $v, w$ , which does not contain  $\begin{pmatrix} 1 & 1 \\ 1 & * \end{pmatrix}$  as a submatrix. We are now done by our comments at the beginning of the proof.  $\square$

Having proved Proposition 4.2.4, we merely need to put the pieces together to finish the proof of our main theorem. This proposition shows that there is a finite 0/1 matrix  $M$  with no submatrix of the form

$$\begin{pmatrix} 1 & 1 \\ 1 & * \end{pmatrix}$$

such that the simple permutations of  $\text{Av}(24153, 31524, \ell \cdots 21)$  are contained in  $\text{Grid}(M)$ . It follows that

$$\text{Av}(24153, 31524, \ell \cdots 21) \subseteq \langle \text{Grid}(M) \rangle.$$

Moreover,  $M$  is a forest because if it were to contain a cycle, it would have to contain a submatrix of the form  $\begin{pmatrix} 1 & 1 \\ 1 & * \end{pmatrix}$ . Therefore the permutation class  $\text{Av}(24153, 31524, \ell \cdots 21)$  is wqo by Theorem 4.1.3.

**Theorem 4.2.2.** *For every  $\ell$ , the permutation class  $\text{Av}(24153, 31524, \ell \cdots 21)$  is wqo. Therefore the class of permutation graphs omitting  $P_5$  and  $K_\ell$  is also wqo.*

### 4.3 Permutation Graphs Omitting $P_6$ , $P_7$ or $P_8$

In this section, we establish the following:

**Proposition 4.3.1.** *The following three classes of graphs are not wqo:*

- (1) *the  $P_6$ ,  $K_6$ -free permutation graphs,*
- (2) *the  $P_7$ ,  $K_5$ -free permutation graphs, and*
- (3) *the  $P_8$ ,  $K_4$ -free permutation graphs.*

In order to prove these three classes are not wqo, it suffices to show that each class contains an infinite antichain. This is done by showing that the related permutation classes contain “generalised” grid classes, for which infinite antichains are already known. In general, we cannot immediately guarantee that the permutation antichain translates to a graph antichain, but we will show that this is in fact the case for the three we construct here.

We must first introduce *generalised* grid classes. Suppose that  $\mathcal{M}$  is a  $t \times u$  matrix of permutation classes (we use calligraphic font for matrices containing permutation classes). An  $\mathcal{M}$ -*gridding* of the permutation  $\pi$  of length  $n$  in this context is a pair of sequences  $1 = c_1 \leq \cdots \leq c_{t+1} = n + 1$  (the column divisions) and  $1 = r_1 \leq \cdots \leq r_{u+1} = n + 1$  (the row divisions) such that for all  $1 \leq k \leq t$  and  $1 \leq \ell \leq u$ , the entries of  $\pi$  from indices  $c_k$  up to but not including  $c_{k+1}$ , which have values from  $r_\ell$  up to but not including  $r_{\ell+1}$  are either empty or order isomorphic to an element of  $\mathcal{M}_{k,\ell}$ . The *grid class* of  $\mathcal{M}$ , written  $\text{Grid}(\mathcal{M})$ , consists of all permutations which possess an  $\mathcal{M}$ -gridding. The notion of monotone griddability can be analogously defined, but we do not require this.

Here, our generalised grid classes are formed from gridding matrices which contain the monotone class  $\text{Av}(21)$ , and a non-monotone permutation class denoted  $\oplus 21$ . This is formed by taking all finite subpermutations of the infinite permutation  $21436587 \cdots$ . In terms of minimal forbidden elements, we have  $\oplus 21 = \text{Av}(321, 231, 312)$ .

Our proof of Proposition 4.3.1 requires some further theory to ensure we can convert the permutation antichains we construct into graph antichains. The primary issue is that a permutation graph  $G$  can have several different corresponding

permutations. With this in mind, let

$$\Pi(G) = \{\text{permutations } \pi : G_\pi \cong G\}$$

denote the set of permutations each of which corresponds to the permutation graph  $G$ .

Denote by  $\pi^{-1}$  the (group-theoretic) *inverse* of  $\pi$ , by  $\pi^{\text{rc}}$  the *reverse-complement* (formed by reversing the order of the entries of  $\pi$ , then replacing each entry  $i$  by  $|\pi| - i + 1$ ), and by  $(\pi^{-1})^{\text{rc}}$  the *inverse-reverse-complement*, formed by composing the two previous operations (in either order, as the two operations commute). It is then easy to see that if  $\pi \in \Pi(G)$ , then  $\Pi(G)$  must also contain all of  $\pi^{-1}$ ,  $\pi^{\text{rc}}$  and  $(\pi^{-1})^{\text{rc}}$ . However, it is possible that  $\Pi(G)$  may contain other permutations, and this depends on the graph-theoretic analogue of the substitution decomposition, which is called the *modular decomposition*.

We also need to introduce the graph analogues of intervals and simplicity, which have different names in that context. A *module*  $M$  in a graph  $G$  is a set of vertices such that for every  $u, v \in M$  and  $w \in V(G) \setminus M$ ,  $u$  is adjacent to  $w$  if and only if  $v$  is adjacent to  $w$ . A graph  $G$  is said to be *prime* if it has no nontrivial modules, that is, any module  $M$  of  $G$  satisfies  $|M| = 0, 1$ , or  $|V(G)|$ .

The following result, arising as a consequence of Gallai's work on transitive orientations, gives us some control over  $\Pi(G)$ :

**Proposition 4.3.2** ([Gallai, 1967]). *If  $G$  is a prime permutation graph, then, up to the symmetries inverse, reverse-complement, and inverse-reverse-complement,  $\Pi(G)$  contains a unique permutation.*

We now extend Proposition 4.3.2 to suit our purposes. Consider a simple permutation  $\sigma$  of length  $m \geq 4$ , and form the permutation  $\pi$  by inflating two<sup>2</sup> of the entries of  $\sigma$  each by the permutation 21. Note that  $G_{21} = K_2$ , and  $\Pi(K_2) = \{21\}$ . In the correspondence between graphs and permutations, modules map to intervals and vice versa, so prime graphs correspond to simple permutations, and there is an analogous result to Proposition 4.1.2 for (permutation) graphs.

Thus, for any permutation  $\rho \in \Pi(G_\pi)$ , it follows that  $\rho$  must be constructed by inflating two entries of some simple permutation  $\tau$  by the permutation 21. Moreover,  $\tau$  must be one of  $\sigma$ ,  $\sigma^{-1}$ ,  $\sigma^{\text{rc}}$  or  $(\sigma^{-1})^{\text{rc}}$ , and the entries of  $\tau$  which are inflated are determined by which entries of  $\sigma$  were inflated, and which of the four symmetries of  $\sigma$  is equal to  $\tau$ . In other words, we still have  $\Pi(G_\pi) = \{\pi, \pi^{-1}, \pi^{\text{rc}}, (\pi^{-1})^{\text{rc}}\}$ .

---

<sup>2</sup>It is, in fact, possible to inflate more than two entries and establish the same result, but we do not require this here.

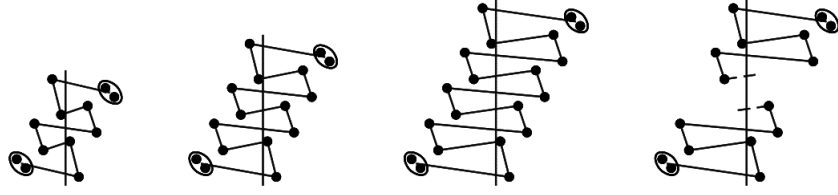


Figure 4.6: The first three elements and the general structure of the infinite antichain  $A$  in  $\text{Grid}(\oplus^{21} \oplus^{21})$

We require one further easy observation:

**Lemma 4.3.1.** *If  $G$  and  $H$  are permutation graphs such that  $H \leq G$ , then for any  $\pi \in \Pi(G)$  there exists  $\sigma \in \Pi(H)$  such that  $\sigma \leq \pi$ .*

*Proof.* Given any  $\pi \in \Pi(G)$ , let  $\sigma$  denote the subpermutation of  $\pi$  formed from the entries of  $\pi$  which correspond to the vertices of an embedding of  $H$  as an induced subgraph of  $G$ . Clearly  $G_\sigma \cong H$ , so  $\sigma \in \Pi(H)$  and  $\sigma \leq \pi$ , as required.  $\square$

We are now in a position to prove Proposition 4.3.1. Since the techniques are broadly similar for all three cases, we will give the details for case (2), and only outline the key steps for the other two cases.

*Proof of Proposition 4.3.1 (2).* First, the graph  $P_7$  corresponds to two permutations, namely 3152746 and 2416375 respectively. Thus the class of  $P_7$ ,  $K_5$ -free permutation graphs corresponds to the permutation class  $\text{Av}(3152746, 2416375, 54321)$ . This permutation class contains the grid class

$$\text{Grid}\left(\oplus^{21} \oplus^{21}\right),$$

because this grid class avoids the permutations 241635 (contained in both 3152746 and 2416375), and 54321.

We now follow the recipe given by Brignall [2012] to construct an infinite antichain which lies in  $\text{Grid}(\oplus^{21} \oplus^{21})$ . Call the resulting antichain  $A$ , the first three elements and general term of which are illustrated in Figure 4.6. This antichain is related to the “parallel” antichain in Murphy’s thesis [Murphy, 2002].

Now set  $G_A = \{G_\pi : \pi \in A\}$ , and note that  $G_A$  is contained in the class of permutation graphs omitting  $P_7$  and  $K_5$  by Lemma 4.3.1. If  $G_A$  is an antichain of graphs, then we are done, so suppose for a contradiction that there exists  $G, H \in G_A$  with  $H \leq G$ . Take the permutation  $\pi \in A$  for which  $G_\pi = G$ . Applying Lemma 4.3.1, there exists  $\sigma \in \Pi(H)$  such that  $\sigma \leq \pi$ . We cannot have  $\sigma \in A$

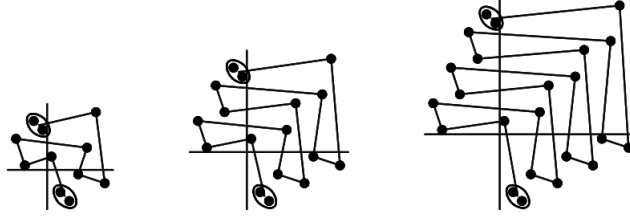


Figure 4.7: The first three elements of an infinite antichain in  $\text{Grid} \left( \begin{smallmatrix} \oplus^{21} \text{Av}(21) \\ \oplus^{21} \end{smallmatrix} \right)$

since  $A$  is an antichain, so  $\sigma$  must be some other permutation with the same graph. Choose  $\tau \in A \cap \Pi(H)$ .

It is easy to check that the only non-trivial intervals in any permutation in  $A$  are the two pairs of points in Figure 4.6 which are circled. Thus,  $\tau$  is formed by inflating two entries of a simple permutation each by a copy of 21. By the comments after Proposition 4.3.2, we conclude that  $\sigma$  is equal to one of  $\tau^{-1}$ ,  $\tau^{\text{rc}}$  or  $(\tau^{-1})^{\text{rc}}$ .

Note that every permutation in  $A$  is closed under reverse-complement, so  $\tau = \tau^{\text{rc}}$  and thus  $\sigma \neq \tau^{\text{rc}}$ . If  $\sigma = \tau^{-1}$ , by inspection every element of  $A$  contains a copy of 24531, which means that  $\sigma$  contains  $(24531)^{-1} = 51423$ . Since  $\sigma \leq \pi$ , it follows that  $\pi$  must also contain a copy of 51423. However this is impossible, because 51423 is not in  $\text{Grid} \left( \begin{smallmatrix} \oplus^{21} & \oplus^{21} \end{smallmatrix} \right)$ .

Thus, we must have  $\sigma = (\tau^{-1})^{\text{rc}}$ , in which case  $\sigma$  must contain a copy of  $((24531)^{\text{rc}})^{-1} = 34251$ . However, this permutation is also not in  $\text{Grid} \left( \begin{smallmatrix} \oplus^{21} & \oplus^{21} \end{smallmatrix} \right)$  so cannot be contained in  $\pi$ . Thus  $\sigma \notin \pi$ , and from this final contradiction we conclude that  $H \not\leq G$ , so  $G_A$  is an infinite antichain of permutation graphs, as required.  $\square$

*Sketch proof of Proposition 4.3.1 (1) and (3).* For (1), the class of  $P_6$ ,  $K_6$ -free permutation graphs corresponds to  $\text{Av}(241635, 315264, 654321)$ . This class contains  $\text{Grid} \left( \begin{smallmatrix} \oplus^{21} \text{Av}(21) \\ \oplus^{21} \end{smallmatrix} \right)$  because this grid class does not contain 23154 (contained in 241635), 31254 (contained in 315264) and 654321.

By [Brignall, 2012], this grid class contains an infinite antichain whose first three elements are illustrated in Figure 4.7. The permutations in this antichain have exactly two proper intervals, indicated by the circled pairs of points in each case, and this means that for any permutation  $\pi$  in this antichain,  $\Pi(G_\pi) = \{\pi, \pi^{-1}, \pi^{\text{rc}}, (\pi^{-1})^{\text{rc}}\}$ .

Following the proof of Proposition 4.3.1 (2), it suffices to show that for any permutations  $\sigma$  and  $\pi$  in the antichain, none of  $\sigma$ ,  $\sigma^{-1}$ ,  $\sigma^{\text{rc}}$ , or  $(\sigma^{-1})^{\text{rc}}$  is contained



in  $\pi$ . This is done by identifying permutations which are not contained in any antichain element  $\pi$ , but which are contained in one of the symmetries  $\sigma^{-1}$ ,  $\sigma^{\text{rc}}$ , or  $(\sigma^{-1})^{\text{rc}}$ . We omit the details.

For (3), The  $P_8$ ,  $K_4$ -free permutation graphs correspond to  $\text{Av}(4321, 24163857, 31527486)$ , and this permutation class contains the grid class  $\text{Grid} \begin{pmatrix} \text{Av}(21) & \text{Av}(21) \\ \text{Av}(21) & \text{Av}(21) \end{pmatrix}$ . In this case, we appeal to Murphy and Vatter [2003] for an infinite antichain, and the same technique used to prove cases (1) and (2) can be applied.  $\square$

## 4.4 Conclusion

As shown in Figure 4.1, there are only three cases remaining: permutation graphs avoiding  $\{P_6, K_5\}$ ,  $\{P_6, K_4\}$  and  $\{P_7, K_4\}$ . Due to the absence of an “obvious” infinite antichain in these cases, we conjecture that they are all wqo. However, all three classes contain (for example) the generalised grid class  $\text{Grid} \left( \bigoplus_{21} \text{Av}(21) \right)$ , so these classes all contain simple permutations which are not monotone griddable. Thus our approach would have to be significantly changed to approach this conjecture.

## Part II

# WQO and Clique-width

## Chapter 5

# WQO, Clique-width and The Speed of Hereditary Properties

Daligault et al. [2010] conjectured that if a class of graphs is well-quasi-ordered by induced subgraph relation, then the clique-width of graphs in the class is bounded by a constant. In [Allen et al., 2009] a certain relation between the clique-width and the number of  $n$ -vertex graphs in the class, also known as the speed of the hereditary class, was revealed. The authors proved that the classes of bounded clique-width have at most factorial speed of growth and that all classes with speed below the Bell number  $B_n$  are of bounded clique-width. Moreover, it is not hard to see and we will prove formally in Section 5.5, that the classes below Bell number are subclasses of  $k$ -uniform graphs, and therefore they are finitely defined and well-quasi-ordered by the induced subgraph relation. This motivates us to have a look at the boundary separating classes above the Bell number and classes below the Bell number.

In Section 5.3, we prove the main result of this Chapter: we provide the description of all minimal classes with speed above the Bell number. In Section 5.4 we provide an algorithm which given a finitely defined class decides whether the speed is above or below the Bell number. In Section 5.5 we show that these results allow us to verify the conjecture of Daligault et al. [2010] and decide well-quasi-ordering for a family of finitely defined classes for which a certain graph parameter, which we call distinguishing number, is bounded.

In the next two sections we review some basic concepts regarding the speed of hereditary properties and prove some preliminary results.

## 5.1 The speed of hereditary graph properties and the Bell number

Given a property  $\mathcal{X}$ , we write  $\mathcal{X}_n$  for the number of graphs in  $\mathcal{X}$  with vertex set  $\{1, 2, \dots, n\}$  (that is, we are counting *labelled* graphs). Following [Balogh et al., 2000], we call  $\mathcal{X}_n$  the *speed* of the property  $\mathcal{X}$ .

The speeds of hereditary properties and their asymptotic structure have been extensively studied, originally in the special case of a single forbidden subgraph [Erdős et al., 1986, 1976, Kolaitis et al., 1987, Prömel and Steger, 1991, 1992, 1993], and more recently in general [Alekseev, 1993, Alon et al., 2011, Balogh et al., 2000, 2001, 2005, Scheinerman and Zito, 1994]. These studies showed, in particular, that there is a certain correlation between the speed of a property  $\mathcal{X}$  and the structure of graphs in  $\mathcal{X}$ , and that the rates of the speed growth constitute discrete layers. The first four lower layers have been distinguished in [Scheinerman and Zito, 1994]: these are constant, polynomial, exponential, and factorial layers. In other words, Scheinerman and Zito [1994] showed that some classes of functions do not appear as the speed of any hereditary property, and that there are discrete jumps, for example, from polynomial to exponential speeds.

One more jump in the speed of hereditary properties was identified in [Balogh et al., 2005] and it separates – within the factorial layer – the properties with speeds strictly below the Bell number  $B_n$  from those whose speed is at least  $B_n$ . With a slight abuse of terminology we will refer to these two families of graph properties as properties below and above the Bell number, respectively.

In the study of the boundary separating the two families, i.e. the minimal classes above the Bell number, [Balogh et al., 2005] distinguishes two cases: the case where a certain parameter associated with each class of graphs is finite and the case where this parameter is infinite. In this work, we call this parameter *distinguishing number*. For the case where the distinguishing number is infinite, [Balogh et al., 2005] provides a complete description of minimal classes, of which there are precisely 13. For the case where the distinguishing number is finite, [Balogh et al., 2005] mentions only one minimal class above the Bell number (linear forests) and leaves the question of characterising other minimal classes open.

To answer this question, in the next section we define several technical tools and obtain some preparatory results.

## 5.2 Preliminaries and preparatory results

### 5.2.1 $(\ell, d)$ -graphs and sparsification

Given a graph  $G$  and two vertex subsets  $U, W \subset V(G)$ , define  $\Delta(U, W) = \max\{|N(u) \cap W|, |N(w) \cap U| : u \in U, w \in W\}$ . With  $\bar{N}(u) = V(G) \setminus (N(u) \cup \{u\})$ , let  $\bar{\Delta}(U, W) = \max\{|\bar{N}(u) \cap W|, |\bar{N}(w) \cap U| : w \in W, u \in U\}$ . Note that  $\Delta(U, U)$  is simply the maximum degree in  $G[U]$ .

**Definition 5.2.1.** Let  $G$  be a graph. A partition  $\pi = \{V_1, V_2, \dots, V_{\ell'}\}$  of  $V(G)$  is an  $(\ell, d)$ -partition if  $\ell' \leq \ell$  and for each pair of not necessarily distinct integers  $i, j \in \{1, 2, \dots, \ell'\}$  either  $\Delta(V_i, V_j) \leq d$  or  $\bar{\Delta}(V_i, V_j) \leq d$ . We call the sets  $V_i$  *bags*. A graph  $G$  is an  $(\ell, d)$ -graph if it admits an  $(\ell, d)$ -partition.

It should be clear that, given an  $(\ell, d)$ -partition  $\{V_1, V_2, \dots, V_{\ell'}\}$  of  $V(G)$ , for each  $x \in V(G)$  and  $i \in \{1, 2, \dots, \ell'\}$  either  $|N(x) \cap V_i| \leq d$  or  $|\bar{N}(x) \cap V_i| \leq d$ . In the former case we say that  $x$  is *d-sparse* with respect to  $V_i$  and in the latter case we say  $x$  is *d-dense* with respect to  $V_i$ . Similarly, if  $\Delta(V_i, V_j) \leq d$ , we say  $V_i$  is *d-sparse* with respect to  $V_j$ , and if  $\bar{\Delta}(V_i, V_j) \leq d$ , we say  $V_i$  is *d-dense* with respect to  $V_j$ . We will also say that the pair  $(V_i, V_j)$  is *d-sparse* or *d-dense*, respectively. Note that if the bags are large enough (i.e.,  $\min\{|V_i|\} > 2d + 1$ ), the terms *d-dense* and *d-sparse* are mutually exclusive.

**Definition 5.2.2.** A *strong  $(\ell, d)$ -partition* is an  $(\ell, d)$ -partition each bag of which contains at least  $5 \times 2^\ell d$  vertices; a *strong  $(\ell, d)$ -graph* is a graph which admits a strong  $(\ell, d)$ -partition.

Given any strong  $(\ell, d)$ -partition  $\pi = \{V_1, V_2, \dots, V_{\ell'}\}$  we define an equivalence relation  $\sim$  on the bags by putting  $V_i \sim V_j$  if and only if for each  $k$ , either  $V_k$  is *d-dense* with respect to both  $V_i$  and  $V_j$ , or  $V_k$  is *d-sparse* with respect to both  $V_i$  and  $V_j$ . Let us call a partition  $\pi$  *prime* if all its  $\sim$ -equivalence classes are of size 1. If the partition  $\pi$  is not prime, let  $p(\pi)$  be the partition consisting of unions of bags in the  $\sim$ -equivalence classes for  $\pi$ .

We proceed to showing that the partition  $p(\pi)$  of a strong  $(\ell, d)$ -graph does not depend on the choice of a strong  $(\ell, d)$ -partition  $\pi$ . The following three lemmas are the ingredients for the proof of this result.

**Lemma 5.2.1.** Consider any strong  $(\ell, d)$ -graph  $G$  with any strong  $(\ell, d)$ -partition  $\pi$ . Then  $p(\pi)$  is an  $(\ell, \ell d)$ -partition with at least  $5 \times 2^\ell d$  vertices in each bag.

*Proof.* Consider two bags  $W_1, W_2 \in p(\pi)$ . By definition  $W_i = \bigcup_{s \in S_i} V_s$  for some  $S_i \subset \{1, 2, \dots, \ell'\}$ ,  $i = 1, 2$ . Also, it is not hard to see that for all  $(s_1, s_2) \in S_1 \times S_2$

the pairs  $(V_{s_1}, V_{s_2})$  are all either  $d$ -dense or  $d$ -sparse. If they are  $d$ -sparse, then for any  $s_1 \in S_1$  we have  $\Delta(V_{s_1}, W_2) \leq \sum_{s_2 \in S_2} \Delta(V_{s_1}, V_{s_2}) \leq |S_2|d$ . Since this holds for every  $s_1 \in S_1$ , for all  $x \in W_1$  we have that  $|N(x) \cap W_2| \leq |S_2|d$ . Similarly we conclude that for all  $x \in W_2$  we have  $|N(x) \cap W_1| \leq |S_1|d$ . Therefore,  $\Delta(W_1, W_2) \leq \max(|S_1|, |S_2|)d \leq \ell d$ . If the pairs of bags are  $d$ -dense, a similar argument proves that  $\overline{\Delta}(W_1, W_2) \leq \ell d$ . Hence the partition  $p(\pi)$  is an  $(\ell, \ell d)$ -partition. As it is obtained by unifying some bags from a strong  $(\ell, d)$ -partition, we conclude that each bag is of size at least  $5 \times 2^\ell d$ .  $\square$

**Lemma 5.2.2** (Lemma 10 of [Balogh et al., 2005]). *Let  $G$  be a graph with an  $(\ell, d)$ -partition  $\pi$ . If two vertices  $x, y \in G$  are in the same bag  $V_k$ , then the symmetric difference of their neighbourhoods  $N(x) \ominus N(y)$  is of size at most  $2\ell d$ .*  $\square$

**Lemma 5.2.3.** *Let  $G$  be a graph with a strong  $(\ell, d)$ -partition  $\pi$ . If two vertices  $x, y \in V(G)$  belong to different bags of the partition  $p(\pi)$ , then the symmetric difference of their neighbourhoods  $N(x) \ominus N(y)$  is of size at least  $5 \times 2^\ell d - 2d$ .*

*Proof.* Take any two vertices  $x \in V_i$  and  $y \in V_j$  with bags  $V_i$  and  $V_j$  belonging to different  $\sim$ -equivalence classes. Then there is a bag  $V_k$  such that one of the pairs  $(V_i, V_k)$  and  $(V_j, V_k)$  is  $d$ -dense and the other one is  $d$ -sparse; without loss of generality, suppose that  $(V_i, V_k)$  is  $d$ -sparse and  $(V_j, V_k)$  is  $d$ -dense. Then, in particular,  $|N(x) \cap V_k| \leq d$  and  $|N(y) \cap V_k| \geq |V_k| - d$ . Hence  $|N(x) \ominus N(y)| \geq |N(y) \setminus N(x)| \geq |V_k| - 2d \geq 5 \times 2^\ell d - 2d$ .  $\square$

We are now ready to prove the uniqueness of  $p(\pi)$ .

**Theorem 5.2.1.** *Let  $G$  be a strong  $(\ell, d)$ -graph with strong  $(\ell, d)$ -partitions  $\pi$  and  $\pi'$ . Then  $p(\pi) = p(\pi')$ .*

*Proof.* Assume two vertices  $x, y \in V(G)$  are in the same bag of the partition  $p(\pi)$ . By Lemma 5.2.1,  $p(\pi)$  is an  $(\ell, \ell d)$ -partition, so applying Lemma 5.2.2 to  $p(\pi)$  we obtain  $|N(x) \ominus N(y)| \leq 2\ell(\ell d) = 2\ell^2 d < 5 \times 2^\ell d - 2d$ . Thus by Lemma 5.2.3,  $x$  and  $y$  are in the same bag of  $\pi'$ , so they are in the same bag of  $p(\pi')$ . Hence, using symmetry,  $x$  and  $y$  are in the same bag of  $p(\pi)$  if and only if they are in the same bag of  $p(\pi')$ . We deduce that the partitions are the same, i.e.,  $p(\pi) = p(\pi')$ .  $\square$

With any strong  $(\ell, d)$ -partition  $\pi = \{V_1, V_2, \dots, V_{\ell'}\}$  of a graph  $G$  we can associate a *density graph* (with loops allowed)  $H = H(G, \pi)$ : the vertex set of  $H$  is  $\{1, 2, \dots, \ell'\}$  and there is an edge joining  $i$  and  $j$  if and only if  $(V_i, V_j)$  is a  $d$ -dense pair (so there is a loop at  $i$  if and only if  $V_i$  is  $d$ -dense).

For a graph  $G$ , a vertex partition  $\pi = \{V_1, V_2, \dots, V_{\ell'}\}$  of  $G$  and a graph with loops allowed  $H$  with vertex set  $\{1, 2, \dots, \ell'\}$ , we define (as in [Balogh et al., 2000]) the  $H, \pi$ -transform  $\psi(G, \pi, H)$  to be the graph obtained from  $G$  by replacing  $G[V_i, V_j]$  with its bipartite complement for every pair  $(V_i, V_j)$  for which  $ij$  is an edge of  $H$ , and replacing  $G[V_i]$  with its complement for every  $V_i$  for which there is a loop at the vertex  $i$  in  $H$ .

Moreover, if  $\pi$  is a strong  $(\ell, d)$ -partition we define  $\phi(G, \pi) = \psi(G, \pi, H(G, \pi))$ . Note that  $\pi$  is a strong  $(\ell, d)$ -partition for  $\phi(G, \pi)$  and each pair  $(V_i, V_j)$  is  $d$ -sparse in  $\phi(G, \pi)$ . We now show that the result of this “sparsification” does not depend on the initial strong  $(\ell, d)$ -partition.

**Proposition 5.2.1.** *Let  $G$  be a strong  $(\ell, d)$ -graph. Then for any two strong  $(\ell, d)$ -partitions  $\pi$  and  $\pi'$ , the graph  $\phi(G, \pi)$  is identical to  $\phi(G, \pi')$ .*

*Proof.* Suppose that  $\pi = \{U_1, U_2, \dots, U_{\ell'}\}$  and  $\pi' = \{V_1, V_2, \dots, V_{\ell'}\}$ . By Theorem 5.2.1,  $p(\pi) = p(\pi') = \{W_1, W_2, \dots, W_{\ell'}\}$ . Consider two vertices  $x, y$  of  $G$ . Let  $i, j, i', j', i'', j''$  be the indices such that  $x \in U_i, x \in V_{i'}, x \in W_{i''}, y \in U_j, y \in V_{j'}, y \in W_{j''}$ . As the partitions have at least  $5 \times 2^\ell d$  vertices in each bag,  $\ell d$ -dense and  $\ell d$ -sparse are mutually exclusive properties. Hence the pair  $(U_i, U_j)$  is  $d$ -sparse if and only if  $(W_{i''}, W_{j''})$  is  $\ell d$ -sparse if and only if  $(V_{i'}, V_{j'})$  is  $d$ -sparse; and analogously for dense pairs. Therefore  $xy$  is an edge of  $\phi(G, \pi)$  if and only if it is an edge of  $\phi(G, \pi')$ .  $\square$

Proposition 5.2.1 motivates the following definition, originating from [Balogh et al., 2000].

**Definition 5.2.3.** For a strong  $(\ell, d)$ -graph  $G$ , its *sparsification* is  $\phi(G) = \phi(G, \pi)$  for any strong  $(\ell, d)$ -partition  $\pi$  of  $G$ .

The next lemma provides a sufficient condition for the sparsification construction to preserve induced subgraph containment.

**Lemma 5.2.4.** *Suppose  $G_1$  is a strong  $(l, d)$ -graph with a strong  $(l, d)$ -partition  $\pi_1$  and  $G_2$  an  $(l, d)$ -graph with an  $(l, d)$ -partition  $\pi_2$ . Let  $f$  be an embedding of  $G_1$  to  $G_2$ . If the number of bags  $l_1$  in  $p(\pi_1) = \{W_1^1, W_2^1, \dots, W_{l_1}^1\}$  is not smaller than the number of bags  $l_2$  in  $\pi_2 = \{W_1^2, W_2^2, \dots, W_{l_2}^2\}$ , then*

- (1)  $l_1 = l_2$  and there is a permutation  $\sigma_f$  on  $[l_1]$  such that  $f(W_i^1) \subseteq W_{\sigma_f(i)}^2$ ;
- (2)  $p(\pi_1)$  is a strong  $(l, d)$ -partition of  $G_1$  and  $\pi_2$  is a strong  $(l, d)$ -partition of  $G_2$ ;
- (3) the pair  $(W_i^1, W_j^1)$  is  $d$ -dense if and only if the pair  $(W_{\sigma_f(i)}, W_{\sigma_f(j)})$  is  $d$ -dense;
- (4)  $xy \in E(\phi(G_1))$  if and only if  $f(x)f(y) \in E(\phi(G_2))$ .

*Proof.* Consider  $G_1$  and  $G_2$  as in the statement and an embedding  $f : G_1 \rightarrow G_2$ . For  $i \in [l_1]$  put  $S(i) = \{x \in [l_2] : f(W_i^1) \cap W_x^2 \neq \emptyset\}$ .

We claim that for  $i \neq j$  we have  $S(i) \cap S(j) = \emptyset$ . Suppose, for the sake of contradiction, that for some  $i, j \in [l_1]$ ,  $i \neq j$  we have  $S(i) \cap S(j) \neq \emptyset$  and let  $x \in S(i) \cap S(j)$ . Then there are two vertices  $v_i \in W_i^1$ ,  $v_j \in W_j^1$  such that  $f(v_i), f(v_j) \in W_x^2$ . Now  $p(\pi_2)$  is an  $(l, ld)$ -partition by Lemma 5.2.1 and  $f(v_i), f(v_j)$  are two vertices in the same bag of  $p(\pi_2)$ . Thus by Lemma 5.2.2 we conclude that  $|N(f(v_i)) \ominus N(f(v_j))| \leq 2l(ld) = 2l^2d$ . As  $f$  is an embedding, we have  $f(N(v_i) \ominus N(v_j)) \subseteq N(f(v_i)) \ominus N(f(v_j))$ ; hence  $|N(v_i) \ominus N(v_j)| \leq 2l^2d$ . However,  $v_i$  and  $v_j$  belong to different bags of  $p(\pi_1)$ , so by Lemma 5.2.3 we obtain that  $|N(v_i) \ominus N(v_j)| \geq 5 \times 2^l d - 2d$ . This implies  $5 \times 2^l d - 2d \leq 2l^2d$  which is a contradiction.

So for all  $i \neq j$  we have  $S(i) \cap S(j) = \emptyset$ . This implies that  $l_2 \geq |S(1) \cup S(2) \cup \dots \cup S(l_1)| = |S(1)| + |S(2)| + \dots + |S(l_1)| \geq l_1$ . By assumption,  $l_1 \geq l_2$ , so  $l_1 = l_2$ . This implies that  $|S(i)| = 1$  for each  $i$ . Therefore there exists a (unique) permutation  $\sigma_f$  of  $[l_1]$  such that  $f(W_i^1) \subseteq W_{\sigma_f(i)}^2$  for each  $i$ , which proves (1).

Now (2) follows easily from (1). Indeed,  $p(\pi_1)$  is a strong  $(l, d)$ -partition because each bag contains at least  $5 \times 2^l d$  vertices and for all  $i, j \in [l_1]$  we have  $\Delta(W_i^1, W_j^1) \leq \Delta(W_{\sigma_f(i)}^2, W_{\sigma_f(j)}^2) \leq d$ . Also, each bag  $W_i^2$  is of size at least  $5 \times 2^l d$  because it contains the image of some bag  $W_j^1$  under an injective mapping  $f$ . This implies that  $\pi_2$  is a strong  $(l, d)$ -partition.

It is not hard to see that  $W_i^1$  is  $d$ -dense with respect to  $W_j^1$  if and only if  $W_{\sigma_f(i)}^2$  is  $d$ -dense with respect to  $W_{\sigma_f(j)}^2$ . Indeed, if one pair is  $d$ -dense and the other  $d$ -sparse, then since  $f(W_i^1) \subseteq W_{\sigma_f(i)}^2$  and  $f(W_j^1) \subseteq W_{\sigma_f(j)}^2$  we have that  $(W_i^1, W_j^1)$  is both  $d$ -sparse and  $d$ -dense. This implies that  $|W_i^1| \leq 2d + 1$ , a contradiction.

Finally, to show (4), consider any  $x, y \in V(G_1)$ ,  $x \in W_i^1$  and  $y \in W_j^1$ . Then by definition  $xy \in E(\phi(G_1))$  if and only if  $xy \in E(G_1)$  and  $(W_i^1, W_j^1)$  is  $d$ -sparse or  $xy \notin E(G_1)$  and  $(W_i^1, W_j^1)$   $d$ -dense. But as  $xy \in E(G_1)$  if and only if  $f(x)f(y) \in E(G_2)$ , and  $(W_i^1, W_j^1)$  is  $d$ -sparse ( $d$ -dense, respectively) if and only if  $(W_{\sigma_f(i)}^2, W_{\sigma_f(j)}^2)$  is  $d$ -sparse ( $d$ -dense, respectively), we conclude that the statement is equivalent to saying that  $f(x)f(y) \in E(\phi(G_2))$ .  $\square$

### 5.2.2 Distinguishing number $k_\chi$

In this section, we discuss the distinguishing number of a hereditary graph property, which is an important parameter introduced by Balogh et al. [2000].

Given a graph  $G$  and a set  $X = \{v_1, \dots, v_t\} \subseteq V(G)$ , we say that the disjoint subsets  $U_1, \dots, U_m$  of  $V(G)$  are *distinguished* by  $X$  if for each  $i$ , all vertices of  $U_i$  have the same neighbourhood in  $X$ , and for each  $i \neq j$ , vertices  $x \in U_i$  and  $y \in U_j$



have different neighbourhoods in  $X$ . We also say that  $X$  *distinguishes* the sets  $U_1, U_2, \dots, U_m$ .

**Definition 5.2.4.** Given a hereditary property  $\mathcal{X}$ , we define the *distinguishing number*  $k_{\mathcal{X}}$  as follows:

- (a) If for all  $k, m \in \mathbb{N}$  we can find a graph  $G \in \mathcal{X}$  that admits some  $X \subset V(G)$  distinguishing at least  $m$  sets, each of size at least  $k$ , then put  $k_{\mathcal{X}} = \infty$ .
- (b) Otherwise, there must exist a pair  $(k, m)$  such that any vertex subset of any graph  $G \in \mathcal{X}$  distinguishes at most  $m$  sets of size at least  $k$ . We define  $k_{\mathcal{X}}$  to be the minimum value of  $k$  in all such pairs.

In [Balogh et al., 2000], the authors show that the speed of any hereditary property  $\mathcal{X}$  with  $k_{\mathcal{X}} = \infty$  is above the Bell number. To study the classes with  $k_{\mathcal{X}} < \infty$ , in the next sections we will use the following results from their paper:

**Lemma 5.2.5** (Lemma 27 of [Balogh et al., 2000]). *If  $\mathcal{X}$  is a hereditary property with finite distinguishing number  $k_{\mathcal{X}}$ , then there exist absolute constants  $\ell_{\mathcal{X}}$ ,  $d_{\mathcal{X}}$  and  $c_{\mathcal{X}}$  such that for all  $G \in \mathcal{X}$ , the graph  $G$  contains an induced subgraph  $G'$  such that  $G'$  is a strong  $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph and  $|V(G) \setminus V(G')| < c_{\mathcal{X}}$ .*  $\square$

**Theorem 5.2.2** (Theorem 28 of [Balogh et al., 2000]). *Let  $\mathcal{X}$  be a hereditary property with  $k_{\mathcal{X}} < \infty$ . Then  $\mathcal{X}_n \geq n^{(1+o(1))n}$  if and only if for every  $m$  there exists a strong  $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph  $G$  in  $\mathcal{X}$  such that its sparsification  $\phi(G)$  has a component of order at least  $m$ .*  $\square$

## 5.3 Structure of minimal classes above Bell

In this section, we describe minimal classes with speed above the Bell number. In [Balogh et al., 2005], the authors characterised all minimal classes with infinite distinguishing number. In Section 5.3.1 we report this result and prove additionally that each of these classes can be characterised by finitely many forbidden induced subgraphs. Then in Section 5.3.2 we move on to the case of finite distinguishing number, which had been left open in [Balogh et al., 2005].

### 5.3.1 Infinite distinguishing number

**Theorem 5.3.1** ([Balogh et al., 2005]). *Let  $\mathcal{X}$  be a hereditary graph property with  $k_{\mathcal{X}} = \infty$ . Then  $\mathcal{X}$  contains at least one of the following (minimal) classes:*

- (a) *the class  $\mathcal{K}_1$  of all graphs each of whose connected components is a clique;*
- (b) *the class  $\mathcal{K}_2$  of all star forests;*

- (c) the class  $\mathcal{K}_3$  of all graphs whose vertex set can be split into an independent set  $I$  and a clique  $Q$  so that every vertex in  $Q$  has at most one neighbour in  $I$ ;
- (d) the class  $\mathcal{K}_4$  of all graphs whose vertex set can be split into an independent set  $I$  and a clique  $Q$  so that every vertex in  $I$  has at most one neighbour in  $Q$ ;
- (e) the class  $\mathcal{K}_5$  of all graphs whose vertex set can be split into two cliques  $Q_1, Q_2$  so that every vertex in  $Q_2$  has at most one neighbour in  $Q_1$ ;
- (f) the class  $\mathcal{K}_6$  of all graphs whose vertex set can be split into two independent sets  $I_1, I_2$  so that the neighbourhoods of the vertices in  $I_1$  are linearly ordered by inclusion (that is, the class of all chain graphs);
- (g) the class  $\mathcal{K}_7$  of all graphs whose vertex set can be split into an independent set  $I$  and a clique  $Q$  so that the neighbourhoods of the vertices in  $I$  are linearly ordered by inclusion (that is, the class of all threshold graphs);
- (h) the class  $\overline{\mathcal{K}}_i$  of all graphs whose complement belongs to  $\mathcal{K}_i$  as above, for some  $i \in \{1, 2, \dots, 6\}$  (note that the complementary class of  $\mathcal{K}_7$  is  $\mathcal{K}_7$  itself).

As an aside, it is perhaps worth noting that each of the minimal classes admits an infinite universal graph. To be specific,  $\mathcal{K}_1$  is the age (the class of all finite induced subgraphs) of  $U_1$ , the disjoint union of  $\omega$  cliques, each of order  $\omega$ . The remaining universal graphs are depicted in Figure 5.1; a grey oval indicates a clique (of order  $\omega$ ).

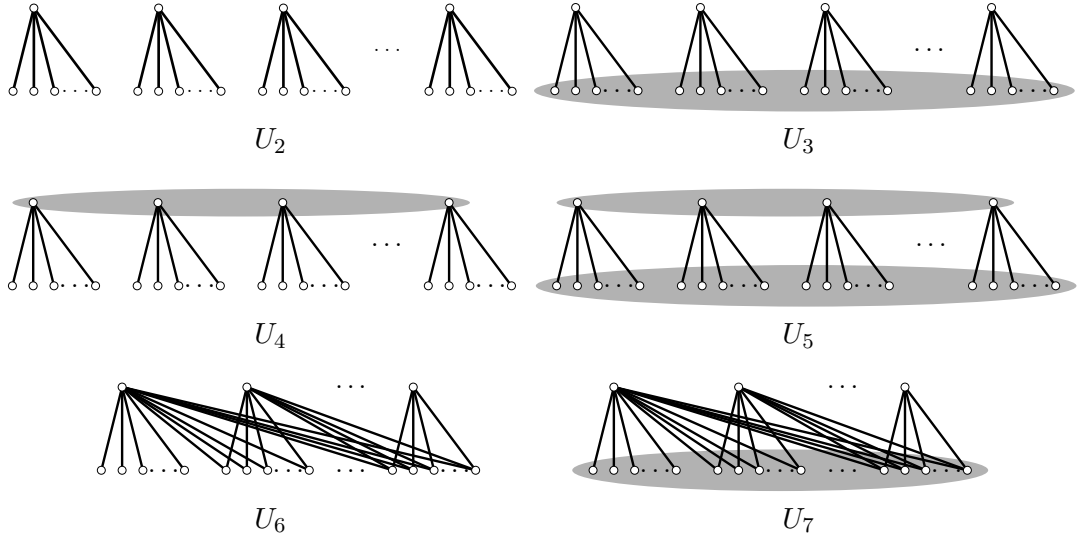


Figure 5.1: The universal graphs

Aiming to prove that each of the classes above is defined by forbidding finitely many induced subgraphs, we first state an older result by Földes and Hammer about split graphs of which we make use in our proof. A *split graph* is a graph whose vertex

set can be split into an independent set and a clique.

**Theorem 5.3.2** ([Földes and Hammer, 1977]). *The class of all split graphs is exactly the class  $\text{Free}(2K_2, C_4, C_5)$ .*

Before showing the characterisation of the classes  $\mathcal{K}_1$ – $\mathcal{K}_6$  in terms of forbidden induced subgraphs, we introduce some of the less commonly appearing graphs: the *claw*  $K_{1,3}$ , the *3-fan*  $F_3$ , the *diamond*  $K_4^-$ , and the graph  $H_6$  (Fig. 5.2).

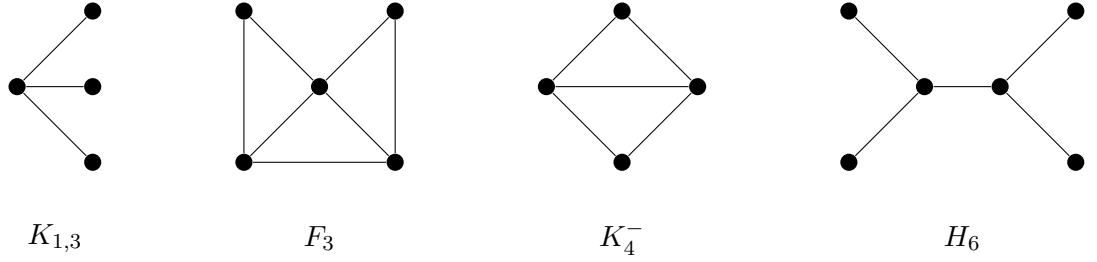


Figure 5.2: Some small graphs

**Theorem 5.3.3.** *Each of the classes of Theorem 5.3.1 is defined by finitely many forbidden induced subgraphs.*

*Proof.* First, observe that if we define  $\overline{\mathcal{X}}$  as the class of the complements of all graphs in  $\mathcal{X}$ , then  $\overline{\text{Free}(\mathcal{F})} = \text{Free}(\overline{\mathcal{F}})$ . Hence if each class  $\mathcal{K}_i$  is defined by finitely many forbidden induced subgraphs, then so is each  $\overline{\mathcal{K}_i}$ .

(a)  $\mathcal{K}_1 = \text{Free}(P_3)$ : It is trivial to check that  $P_3$  does not belong to  $\mathcal{K}_1$ , and any graph not containing an induced  $P_3$  must be a collection of cliques.

(b)  $\mathcal{K}_2 = \text{Free}(K_3, P_4, C_4)$ : Obviously, none of the graphs  $K_3, P_4, C_4$  belongs to  $\mathcal{K}_2$ . Let  $G \in \text{Free}(K_3, P_4, C_4)$ . Since every cycle of length at least 5 contains  $P_4$ ,  $G$  does not contain any cycles; thus  $G$  is a forest. The absence of a  $P_4$  implies that the diameter of any component of  $G$  is at most 2, hence  $G$  is a star forest.

(c)  $\mathcal{K}_3 = \text{Free}(\mathcal{F})$  for  $\mathcal{F} = \{2K_2, C_4, C_5, K_{1,3}, F_3\}$ : It is easy to check that none of the forbidden graphs belong to  $\mathcal{K}_3$ . Let  $G \in \text{Free}(\mathcal{F})$ . By Theorem 5.3.2,  $G$  is a split graph. Split  $G$  into a maximal clique  $Q$  and an independent set  $I$ . Suppose, for the sake of contradiction, that  $Q$  contains a vertex  $u$  with two neighbours  $a, b \in I$ . As we took  $Q$  to be a maximal clique,  $a$  has a non-neighbour  $v$  and  $b$  has a non-neighbour  $w$  in  $Q$ . If  $a, w$  are not adjacent, then the vertices  $a, b, u, w$  induce a claw in  $G$ ; if  $b, v$  are not adjacent, then the vertices  $a, b, u, v$  induce a claw in  $G$ ; otherwise the vertices  $a, b, u, v, w$  induce a 3-fan in  $G$ . In either case we get a contradiction.

(d)  $\mathcal{K}_4 = \text{Free}(\mathcal{F})$  for  $\mathcal{F} = \{2K_2, C_4, C_5, K_4^-\}$ : Again, it is easy to check that none of the forbidden graphs belong to  $\mathcal{K}_4$ . Let  $G \in \text{Free}(\mathcal{F})$ . By Theorem 5.3.2,  $G$  is a split graph. Just like before, split  $G$  into a maximal clique  $Q$  and an independent set  $I$ . Suppose that some vertex  $u$  in  $I$  has two neighbours  $a, b$  in  $Q$ . By maximality of  $Q$ ,  $u$  also has a non-neighbour  $c$  in  $Q$ . But then the vertices  $a, b, c, u$  induce a  $K_4^-$  in  $G$ , a contradiction.

(e) The class  $\overline{\mathcal{K}_5}$  of the complements of the graphs in  $\mathcal{K}_5$  is characterised as the class of all (bipartite) graphs whose vertex set can be split into independent sets  $I_1, I_2$  so that each vertex in  $I_2$  has at most one non-neighbour in  $I_1$ . We show that  $\overline{\mathcal{K}_5} = \text{Free}(\mathcal{F})$  for  $\mathcal{F} = \{K_3, C_5, P_4 + K_1, 2K_2 + K_1, C_4 + K_2, C_4 + 2K_1, H_6\}$ . The reader will kindly check that indeed no graph in  $\mathcal{F}$  belongs to  $\overline{\mathcal{K}_5}$ .

Consider some  $G \in \text{Free}(\mathcal{F})$ ; we will show that  $G \in \overline{\mathcal{K}_5}$ . Observe that  $\mathcal{F}$  prevents  $G$  from having an odd cycle, thus  $G$  is bipartite. We distinguish three cases depending on the structure of the connected components of  $G$ .

First, suppose that  $G$  has at least two non-trivial connected components (that is, connected components that are not just isolated vertices). Because  $G$  is  $(2K_2 + K_1)$ -free, it only has two connected components in all. Being  $C_4$ - and  $P_4$ -free, each component is necessarily a star. Observe that any graph consisting of one or two stars belongs to  $\overline{\mathcal{K}_5}$ .

Next assume that  $G$  has only one non-trivial component and some isolated vertices. The non-trivial component is bipartite and  $P_4$ -free, so it is a biclique. If this biclique contains  $C_4$ , then  $G$  only contains one other isolated vertex; any graph consisting of a biclique and one isolated vertex is in  $\overline{\mathcal{K}_5}$ . Otherwise the biclique is a star; any graph consisting of a star and one or more isolated vertices belongs to  $\overline{\mathcal{K}_5}$ .

Finally, consider  $G$  that is connected. We will show that for any two vertices of  $G$  in different parts, one of them must have at most one non-neighbour in the opposite part. Suppose this is not true and there are  $x, y \in V(G)$  in different parts such that both  $x$  and  $y$  have at least two non-neighbours in the opposite part. Assume first that  $x$  and  $y$  are adjacent. Let  $a$  and  $b$  be two non-neighbours of  $x$ , and let  $c$  and  $d$  be two non-neighbours of  $y$ . Then the graph induced by  $a, b, c$  and  $d$  cannot be a  $C_4$ ,  $P_4$ ,  $P_3 + K_1$ ,  $2K_2$  or  $K_2 + 2K_1$ , because  $G$  is  $(P_4 + K_1, 2K_2 + K_1, C_4 + K_2)$ -free. Hence  $a, b, c$  and  $d$  must induce a  $4K_1$ . As  $G$  is connected,  $a$  must have a neighbour, say  $w$ . However, the vertices  $x, y, a, c$  and  $w$  induce a  $P_4 + K_1$  if  $y$  and  $w$  are adjacent and they induce a  $2K_2 + K_1$  if  $y$  and  $w$  are not adjacent. Therefore,  $x$  and  $y$  must be non-neighbours.

By assumption,  $x$  has another non-neighbour  $a \neq y$  in the opposite part, and  $y$  has another non-neighbour  $b \neq x$  in the opposite part. As  $G$  is connected,  $x$

must have a neighbour, say  $u$ . If  $a$  and  $b$  are adjacent, then  $x, y, u, a$  and  $b$  induce a  $2K_2 + K_1$  if  $u$  is not adjacent to  $b$ , and they induce a  $P_4 + K_1$  if  $u$  is adjacent to  $b$ . Both cases lead to a contradiction as  $G$  is  $(P_4 + K_1, 2K_2 + K_1)$ -free, hence  $a$  and  $b$  cannot be adjacent. Now, as  $G$  is connected,  $y$  must also have a neighbour, say  $v$ . If  $u$  is not adjacent to  $b$ , then  $x, y, u, v$  and  $b$  induce either a  $2K_2 + K_1$  or a  $P_4 + K_1$ , hence  $u$  and  $b$  must be adjacent. By symmetry,  $v$  is adjacent to  $a$ . Now  $u$  and  $v$  must be non-adjacent otherwise  $x, y, u, v, a$  and  $b$  induce an  $H_6$ .

This argument shows that any neighbour of  $x$  must also be a neighbour of  $b$ , any neighbour of  $y$  must also be a neighbour of  $a$ , and that any neighbour of  $x$  cannot be adjacent to any neighbour of  $y$ . This means that the shortest induced path between  $x$  and  $y$  must contain a  $P_6$ , which is a contradiction as  $G$  is  $(P_4 + K_1)$ -free. Therefore, either  $x$  or  $y$  must have at most one non-neighbour. This implies that  $G$  can be split into two independent sets  $I_1, I_2$  such that every vertex in  $I_2$  has at most one neighbour in  $I_1$ , so  $G$  belongs to  $\overline{\mathcal{K}}_5$ .

(f) Chain graphs are characterised by finitely many forbidden induced subgraphs by a result of Yannakakis [1982]; namely,  $\mathcal{K}_6 = \text{Free}(2K_2, K_3, C_5)$ .

(g) Threshold graphs are characterised by finitely many forbidden induced subgraphs by a result of Chvátal and Hammer [1977]; namely,  $\mathcal{K}_7 = \text{Free}(2K_2, P_4, C_4)$ .  $\square$

### 5.3.2 Finite distinguishing number

In this section we provide a characterisation of the minimal classes for the case of finite distinguishing number  $k_{\mathcal{X}}$ . It turns out that these minimal classes consist of  $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graphs, that is, the vertex set of each graph is partitioned into at most  $\ell_{\mathcal{X}}$  bags and dense pairs are defined by a density graph  $H$  (see Lemma 5.2.5). The condition of Theorem 5.2.2 is enforced by long paths (indeed, an infinite path in the infinite universal graph). Thus actually  $d_{\mathcal{X}} \leq 2$  for the minimal classes  $\mathcal{X}$ .

Let  $A$  be a finite alphabet. A *word* is a mapping  $w : S \rightarrow A$ , where  $S = \{1, 2, \dots, n\}$  for some  $n \in \mathbb{N}$  or  $S = \mathbb{N}$ ;  $|S|$  is the *length* of  $w$ , denoted by  $|w|$ . We write  $w_i$  for  $w(i)$ , and we often use the notation  $w = w_1 w_2 w_3 \dots w_n$  or  $w = w_1 w_2 w_3 \dots$ . For  $n \leq m$  and  $w = w_1 w_2 \dots w_n$ ,  $w' = w'_1 w'_2 \dots w'_m$  (or  $w' = w'_1 w'_2 \dots$ ), we say that  $w$  is a *factor* of  $w'$  if there exists a non-negative integer  $s$  such that  $w_i = w'_{i+s}$  for  $1 \leq i \leq n$ ;  $w$  is an *initial segment* of  $w'$  if we can take  $s = 0$ .

Let  $H$  be an undirected graph with loops allowed and with vertex set  $V(H) = A$ , and let  $w$  be a (finite or infinite) word over the alphabet  $A$ . For any increasing sequence  $u_1 < u_2 < \dots < u_m$  of positive integers such that  $u_m \leq |w|$ , define

$G_{w,H}(u_1, u_2, \dots, u_m)$  to be the graph with vertex set  $\{u_1, u_2, \dots, u_m\}$  and an edge between  $u_i$  and  $u_j$  if and only if

- either  $|u_i - u_j| = 1$  and  $w_{u_i}w_{u_j} \notin E(H)$ ,
- or  $|u_i - u_j| > 1$  and  $w_{u_i}w_{u_j} \in E(H)$ .

Let  $G = G_{w,H}(u_1, u_2, \dots, u_m)$  and define  $V_a = \{u_i \in V(G) : w_{u_i} = a\}$  for any  $a \in A$ . Then  $\pi = \pi_w(G) = \{V_a : a \in A\}$  is an  $(|A|, 2)$ -partition, and so  $G$  is an  $(|A|, 2)$ -graph. Moreover,  $\psi(G, \pi, H)$  is a linear forest whose paths are formed by the consecutive segments of integers within the set  $\{u_1, u_2, \dots, u_m\}$ . This partition  $\pi_w(G)$  is called the *letter partition* of  $G$ .

**Definition 5.3.1.** Let  $H$  be an undirected graph with loops allowed and with vertex set  $V(H) = A$ , and let  $w$  be an infinite word over the alphabet  $A$ . Define  $\mathcal{P}(w, H)$  to be the hereditary class consisting of the graphs  $G_{w,H}(u_1, u_2, \dots, u_m)$  for all finite increasing sequences  $u_1 < u_2 < \dots < u_m$  of positive integers.

As we shall see later, all classes  $\mathcal{P}(w, H)$  are above the Bell number. More importantly, all minimal classes above the Bell number have the form  $\mathcal{P}(w, H)$  for some  $w$  and  $H$ . Our goal here is firstly to describe sufficient conditions on the word  $w$  under which  $\mathcal{P}(w, H)$  is a minimal class above the Bell number; moreover, we aim to prove that any hereditary class above the Bell number with finite distinguishing number contains the class  $\mathcal{P}(w, H)$  for some word  $w$  and graph  $H$ . We start by showing that these classes indeed have finite distinguishing number.

**Lemma 5.3.1.** *For any word  $w$  and graph  $H$  with loops allowed, the class  $\mathcal{X} = \mathcal{P}(w, H)$  has finite distinguishing number.*

*Proof.* Put  $\ell = |H|$  and let  $G$  be a graph in  $\mathcal{X}$ . Consider the letter partition  $\pi = \pi_w(G) = \{V_a : a \in V(H)\}$  of  $G$ , which is an  $(\ell, 2)$ -partition. Choose an arbitrary set of vertices  $X \subseteq V(G)$  and let  $\{U_1, U_2, \dots, U_k\}$  be the sets distinguished by  $X$ . If there are subsets  $U_i, U_j$  and  $V_a$  such that  $|V_a \cap U_i| \geq 3$  and  $|V_a \cap U_j| \geq 3$ , then some vertex of  $X$  has at least three neighbours and at least three non-neighbours in  $V_a$ , which contradicts the fact that  $\pi$  is an  $(\ell, 2)$ -partition. Therefore, in the partition  $\{V_a \cap U_i : a \in V(H), 1 \leq i \leq k\}$  we have at most  $\ell$  sets of size at least 3. Note that every set  $U_i$  of size at least  $2\ell + 1$  must contain at least one such set. Hence the family  $\{U_1, U_2, \dots, U_k\}$  contains at most  $\ell$  sets of size at least  $2\ell + 1$ . Since the set  $X$  was chosen arbitrarily, we conclude that  $k_{\mathcal{X}} < 2\ell + 1$ , as required.  $\square$

The graphs  $G_{w,H}(u_1, u_2, \dots, u_n)$  defined on a sequence of consecutive integers will play a special role in our considerations.

**Definition 5.3.2.** If  $u_1, u_2, \dots, u_m$  is a sequence of consecutive integers (i.e.,  $u_{k+1} = u_k + 1$  for each  $k$ ), we call the graph  $G_{w,H}(u_1, u_2, \dots, u_m)$  an  $|H|$ -factor. Notice that each  $|H|$ -factor is an  $(|H|, 2)$ -graph; if its letter partition is a strong  $(|H|, 2)$ -partition, we call it a *strong  $|H|$ -factor*.

Note that if  $G = G_{w,H}(u_1, u_2, \dots, u_m)$  is a strong  $\ell$ -factor, then its sparsification  $\phi(G) = \psi(G, \pi_w(G), H)$  is an induced path with  $m$  vertices.

**Proposition 5.3.1.** *If  $w$  is an infinite word over a finite alphabet  $A$  and  $H$  is a graph on  $A$ , with loops allowed, then the class  $\mathcal{P}(w, H)$  is above the Bell number.*

*Proof.* We may assume that every letter of  $A$  appears in  $w$  infinitely many times: otherwise we can remove a sufficiently long starting segment of  $w$  to obtain a word  $w'$  satisfying this condition, replace  $H$  with its induced subgraph  $H'$  on the alphabet  $A'$  of  $w'$ , and obtain a subclass  $\mathcal{P}(w', H')$  of  $\mathcal{P}(w, H)$  with that property. For sufficiently large  $k$ , the  $|A|$ -factor  $G_k = G_{w,H}(1, \dots, k)$  is a strong  $|A|$ -factor; thus  $\phi(G_k)$  is an induced path of length  $k - 1$ . Having a finite distinguishing number by Lemma 5.3.1, the class  $\mathcal{P}(w, H)$  is above the Bell number by Theorem 5.2.2.  $\square$

**Definition 5.3.3.** An infinite word  $w$  is called *almost periodic* if for any factor  $f$  of  $w$  there is a constant  $k_f$  such that any factor of  $w$  of length at least  $k_f$  contains  $f$  as a factor.

The notion of an almost periodic word plays a crucial role in our characterisation of minimal classes above the Bell number. First, let us show that if  $w$  is almost periodic, then  $\mathcal{P}(w, H)$  is a minimal property above the Bell number. To prove this, we need an auxiliary lemma.

**Lemma 5.3.2.** *Consider  $G = G_{w,H}(u_1, \dots, u_n)$ . If  $G$  is a strong  $(\ell, d)$ -graph and  $\phi(G)$  contains a connected component  $C$  such that  $|C| \geq (2d'^2\ell^2|H|^2 + 1)(m - 1) + 1$ , where  $d' = \max\{d, 2\}$ , then  $V(C)$  contains a sequence of  $m$  consecutive integers.*

*Proof.* Let  $\pi = \{U_1, U_2, \dots, U_{\ell'}\}$  be a strong  $(\ell, d)$ -partition of  $G$ , so that  $\ell' \leq \ell$  and  $\phi(G) = \phi(G, \pi)$ ; let  $\pi' = \{V_a : a \in V(H)\}$  be the letter partition of  $G$ , given by  $V_a = \{u_j \in V(G) : w_{u_j} = a\}$ . Put  $k = |H|$ . Note that  $\pi'$  is an  $(k, 2)$ -partition, hence also an  $(k, d')$ -partition.

Consider the partition  $\rho = \{U_i \cap V_a : 1 \leq i \leq \ell', a \in V(H)\}$  of  $G$ , which is an  $(\ell'k, d')$ -partition. Let  $(U_i \cap V_a, U_j \cap V_b)$  be a pair of non-empty sets such that  $(U_i, U_j)$  is  $d'$ -sparse but  $(V_a, V_b)$  is  $d'$ -dense. Each such pair is both  $d'$ -sparse and  $d'$ -dense, and consequently we have  $|U_i \cap V_a| \leq 2d'$  and  $|U_j \cap V_b| \leq 2d'$ . Moreover, there are at most  $2d'^2$  edges between  $U_i \cap V_a$  and  $U_j \cap V_b$ . Similarly, for any pair

$(U_i \cap V_a, U_j \cap V_b)$  where  $(U_i, U_j)$  is  $d'$ -dense but  $(V_a, V_b)$  is  $d'$ -sparse, we can show that there are at most  $2d'^2$  non-edges between  $U_i \cap V_a$  and  $U_j \cap V_b$ .

Now, the edges of  $\phi(G)$  that are not edges of  $\psi(G, \pi', H)$  are exactly (a) the edges between  $U_i \cap V_a$  and  $U_j \cap V_b$  where  $(U_i, U_j)$  is  $d'$ -sparse and  $(V_a, V_b)$  is  $d'$ -dense, and (b) the non-edges between  $U_i \cap V_a$  and  $U_j \cap V_b$  where  $(U_i, U_j)$  is  $d'$ -dense and  $(V_a, V_b)$  is  $d'$ -sparse. Thus we can bound the number of such edges:  $|E(\phi(G)) \setminus E(\psi(G, \pi', H))| \leq 2d'^2(\ell'k)^2$ . Any other edge of  $\phi(G)$  joins two consecutive integers. Hence any connected component of  $\phi(G)$  that does not contain a sequence of  $m$  consecutive integers consists of at most  $2d'^2(\ell'k)^2 + 1$  blocks of consecutive integers, each of length at most  $m - 1$ ; it can therefore contain at most  $(2d'^2(\ell'k)^2 + 1)(m - 1) \leq (2d'^2\ell^2|H|^2 + 1)(m - 1)$  vertices.  $\square$

**Theorem 5.3.4.** *If  $w$  is an almost periodic infinite word and  $H$  is a finite graph with loops allowed, then  $\mathcal{P}(w, H)$  is a minimal hereditary property above the Bell number.*

*Proof.* The class  $\mathcal{P} = \mathcal{P}(w, H)$  is above the Bell number by Proposition 5.3.1. Thus we only need to show that any proper hereditary subclass  $\mathcal{X}$  of  $\mathcal{P}$  is below the Bell number. Suppose  $\mathcal{X} \subset \mathcal{P}$  and let  $F \in \mathcal{P} \setminus \mathcal{X}$ . By definition of  $\mathcal{P}(w, H)$ , the graph  $F$  is of the form  $G_{w,H}(u_1, \dots, u_n)$  for some positive integers  $u_1 < \dots < u_n$ . Let  $w'$  be the finite word  $w' = w_{u_1}w_{u_1+1}w_{u_1+2} \dots w_{u_n-1}w_{u_n}$ . As  $w$  is almost periodic, there is an integer  $m$  such that any factor of  $w$  of length  $m$  contains  $w'$  as factor. Assume, for the sake of contradiction, that  $\mathcal{X}$  is hereditary and above the Bell number. By Lemma 5.3.1, the distinguishing number of  $\mathcal{P}$ , and hence of  $\mathcal{X}$ , is finite, and therefore, by Lemma 5.2.5 and Theorem 5.2.2, there exists a strong  $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph  $G = G_{w,H}(u'_1, u'_2, \dots, u'_{n'}) \in \mathcal{X}$  such that  $\phi(G)$  has a connected component  $C$  of order at least  $(2d^2\ell^2|H|^2 + 1)(m - 1) + 1$ , where  $\ell = \ell_{\mathcal{X}}$  and  $d = \max\{d_{\mathcal{X}}, 2\}$ . By Lemma 5.3.2, the vertices of  $C$  contain a sequence of  $m$  consecutive integers, i.e.,  $V(G) \supseteq V(C) \supseteq \{u', u'+1, \dots, u'+m-1\}$ . However, the word  $w_{u'}w_{u'+1} \dots w_{u'+m-1}$  contains  $w'$ ; therefore  $G$  contains  $F$ , a contradiction.  $\square$

The existence of minimal classes does not necessarily imply that every class above the Bell number contains a minimal one. However, in our case this is a necessity, as we proceed to show next. Moreover, this will also imply that the minimal classes described in Theorem 5.3.4 are the *only* minimal classes above the Bell number with  $k_{\mathcal{X}}$  finite. To prove this, we first show in the next two lemmas that any class  $\mathcal{X}$  above the Bell number with  $k_{\mathcal{X}}$  finite contains arbitrarily large strong  $\ell_{\mathcal{X}}$ -factors.



**Lemma 5.3.3.** *Let  $\mathcal{X}$  be a hereditary class with speed above the Bell number and with finite distinguishing number  $k_{\mathcal{X}}$ . Then for each  $m$ , the class  $\mathcal{X}$  contains an  $\ell_{\mathcal{X}}$ -factor of order  $m$ .*

*Proof.* From Theorem 5.2.2 it follows that for each  $m$  there is a graph  $G_m \in \mathcal{X}$  which admits a strong  $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -partition  $\{V_1, V_2, \dots, V_{\ell_m}\}$  with  $\ell_m \leq \ell_{\mathcal{X}}$  such that the sparsification  $\phi(G_m)$  has a connected component  $C_m$  of order at least  $(\ell_{\mathcal{X}} d_{\mathcal{X}})^m$ . Fix an arbitrary vertex  $v$  of  $C_m$ . As  $C_m$  is an induced subgraph of  $\phi(G_m)$ , the maximum degree in  $C_m$  is bounded by  $d = \ell_{\mathcal{X}} d_{\mathcal{X}}$ . Hence for any  $k > 0$ , in  $C_m$  there are at most  $d(d-1)^{k-1}$  vertices at distance  $k$  from  $v$ ; so there are at most  $1 + \sum_{k=1}^{m-2} d(d-1)^{k-1} < d^m$  vertices at distance at most  $m-2$  from  $v$ . As  $C_m$  has order at least  $d^m$ , there exists a vertex  $v'$  of distance  $m-1$  from  $v$ . Therefore  $C_m$  contains an induced path  $v = v_1, v_2, \dots, v_m = v'$  of length  $m-1$ . Let  $A = \{1, 2, \dots, \ell_m\}$  and let  $H$  be the graph with vertex set  $A$  and edge between  $i$  and  $j$  if and only if  $V_i$  is  $d_{\mathcal{X}}$ -dense with respect to  $V_j$ . Let  $w_i \in A$  be such that  $v_i \in V_{w_i}$  and define the word  $w = w_1 w_2 \dots w_m$ . The induced subgraph  $G_m[v_1, v_2, \dots, v_m] \cong G_{w, H}(1, 2, \dots, m)$  is an  $\ell_{\mathcal{X}}$ -factor of order  $m$  contained in  $\mathcal{X}$ .  $\square$

**Lemma 5.3.4.** *Let  $\ell$  and  $B$  be positive integers such that  $B \geq 5 \times 2^{\ell+1}$ . Then any  $\ell$ -factor  $G_{w, H}(1, 2, \dots, |w|)$  of order at least  $B^{\ell}$  contains a strong  $\ell$ -factor  $G_{w', H}(1, 2, \dots, |w'|)$  of order at least  $B$  such that  $w'$  is a factor of  $w$ .*

*Proof.* We will prove by induction on  $r \in \{1, 2, \dots, \ell\}$  that any  $\ell$ -factor  $G_{w, H}(1, 2, \dots, B^r)$  on  $B^r$  vertices with at most  $r$  bags in the letter partition contains a strong  $\ell$ -factor on at least  $B$  vertices. For  $r = 1$  the statement holds because any  $\ell$ -factor with one bag in the letter partition of order  $B \geq 5 \times 2^{\ell+1}$  is a strong  $\ell$ -factor. Suppose  $1 < r \leq \ell$ . Then either each letter of  $w = w_1 w_2 \dots w_{B^r}$  appears at least  $B$  times, in which case we are done, or there is a letter  $a = w_i$  which appears less than  $B$  times in  $w$ . Consider the maximal factors of  $w$  that do not contain the letter  $a$ . There are at most  $B$  such factors of  $w$  and the sum of their orders is at least  $B^r - B + 1$ . By the pigeonhole principle, one of these factors has order at least  $B^{r-1}$ ; call this factor  $w''$ . Now  $w''$  contains at most  $r-1$  different letters; thus  $G'' = G_{w'', H}(1, 2, \dots, |w''|)$  is an  $\ell$ -factor of order at least  $B^{r-1}$  for which the letter partition has at most  $(r-1)$  bags. By induction,  $G''$  contains a strong  $\ell$ -factor  $G_{w', H}(1, 2, \dots, |w'|)$  of order at least  $B$  such that  $w'$  is a factor of  $w''$  which is a factor of  $w$ . Hence  $w'$  is a factor of  $w$  and we are done.  $\square$

**Theorem 5.3.5.** *Suppose  $\mathcal{X}$  is a hereditary class above the Bell number with  $k_{\mathcal{X}}$  finite. Then  $\mathcal{X} \supseteq \mathcal{P}(w, H)$  for an infinite almost periodic word  $w$  and a graph  $H$  of order at most  $\ell_{\mathcal{X}}$  with loops allowed.*

*Proof.* From Lemmas 5.3.3 and 5.3.4 it follows that each class  $\mathcal{X}$  with speed above the Bell number with finite distinguishing number  $k_{\mathcal{X}}$  contains an infinite set  $\mathcal{S}$  of strong  $\ell_{\mathcal{X}}$ -factors of increasing order. For each  $H$  on  $\{1, 2, \dots, \ell\}$  with  $1 \leq \ell \leq \ell_{\mathcal{X}}$ , let  $\mathcal{S}_H = \{G_{w,H}(1, \dots, m) \in \mathcal{S}\}$  be the set of all  $\ell_{\mathcal{X}}$ -factors in  $\mathcal{S}$  whose adjacencies are defined using the density graph  $H$ . Then for some (at least one) fixed graph  $H_0$  the set  $\mathcal{S}_{H_0}$  is infinite. Hence also  $L = \{w : G_{w,H_0}(1, \dots, m) \in \mathcal{X}\}$  is an infinite language. As  $\mathcal{X}$  is a hereditary class, the language  $L$  is closed under taking word factors (it is a *factorial language*).

It is not hard to see that any infinite factorial language contains an inclusion-minimal infinite factorial language. So let  $L' \subseteq L$  be a minimal infinite factorial language. It follows from the minimality that  $L'$  is well quasi-ordered by the factor relation, because otherwise removing one word from any infinite antichain and taking all factors of the remaining words would generate an infinite factorial language strictly contained in  $L'$ . Thus there exists an infinite chain  $w^{(1)}, w^{(2)}, \dots$  of words in  $L'$  such that for any  $i < j$ , the word  $w^{(i)}$  is a factor of  $w^{(j)}$ . More precisely, for each  $i$  there is a non-negative integer  $s_i$  such that  $w_k^{(i)} = w_{k+s_i}^{(i+1)}$ . Let  $g(i, k) = k + \sum_{j=1}^{i-1} s_j$ . Now we can define an infinite word  $w$  by putting  $w_k = w_{g(i,k)}^{(i)}$  for the least value of  $i$  for which the right-hand side is defined. (Without loss of generality we get that  $w$  is indeed an infinite word; otherwise we would need to take the reversals of all the words  $w^{(i)}$ .)

Observe that any factor of  $w$  is a factor of some  $w^{(i)}$  and hence in the language  $L'$ . If  $w$  is not almost periodic, then there exists a factor  $f$  of  $w$  such that there are arbitrarily long factors  $f'$  of  $w$  not containing  $f$ . These factors  $f'$  generate an infinite factorial language  $L'' \subset L'$  which does not contain  $f \in L'$ . This contradicts the minimality of  $L'$  and proves that  $w$  is almost periodic.

Because any factor of  $w$  is in  $L$ , any  $G_{w,H_0}(u_1, \dots, u_m)$  is an induced subgraph of some  $\ell_{\mathcal{X}}$ -factor in  $\mathcal{X}$ . Therefore  $\mathcal{P}(w, H_0) \subseteq \mathcal{X}$ .  $\square$

Combining Theorems 5.3.4 and 5.3.5 we derive the main result of this section.

**Corollary 5.3.1.** *Let  $\mathcal{X}$  be a class of graphs with  $k_{\mathcal{X}} < \infty$ . Then  $\mathcal{X}$  is a minimal hereditary class above the Bell number if and only if there exists a finite graph  $H$  with loops allowed and an infinite almost periodic word  $w$  over  $V(H)$  such that  $\mathcal{X} = \mathcal{P}(w, H)$ .*  $\square$

Note that – similarly to the case of infinite distinguishing number – each of the minimal classes  $\mathcal{P}(w, H)$  has an infinite universal graph:  $G_{w,H}(1, 2, 3, \dots)$ . We also remark that a class of graphs, which can be described as  $\mathcal{P}(w, H)$  for some word  $w$  and prime graph  $H$ , would generally admit many different descriptions of

this form. Thus we further investigate the conditions on the word  $w$ , under which  $P(w, H)$  is a minimal class above the Bell number. We have already shown that the word being almost periodic is a sufficient condition. Now we will show a necessary condition: if  $P(w, H)$  is a minimal class above the Bell number and  $H$  is prime then  $w$  is ultimately almost periodic, i.e. it becomes periodic after removing an initial segment of  $w$ . Moreover, we remark that there are examples of minimal classes above the Bell number of the form  $P(w, H)$  where  $H$  is prime and  $w$  is ultimately almost periodic but not almost periodic. Before proving the necessary condition, we need some preparatory results examining embeddings of strong  $l$ -factors.

Given an alphabet  $A$  and a graph  $H$ , there will in general be several ways to express  $G$  as  $G_{w,H}(1, \dots, m)$ . The reversal  $r(w)$  of a word  $w$  produces the same result as  $w$ :  $G_{w,H}(1, \dots, m) \cong G_{r(w),H}(1, \dots, m)$ . Moreover, we can rename the letters of  $w$  according to some automorphism of  $H$ , that is, if  $\theta \in \text{Aut}(H)$ , then  $G_{w,H}(1, \dots, m) = G_{\theta(w),H}(1, \dots, m)$ . Here  $\theta$  acts on the words over  $A$  letter-wise. Hence, with  $R = \{\text{id}, r\}$ , if  $\theta \in \text{Aut}(H) \times R$ , we have  $G_{w,H}(1, \dots, m) \cong G_{\theta(w),H}(1, \dots, m)$ .

The following lemma shows that a (graph) embedding of a strong  $l$ -factor  $G_1$  into another  $l$ -factor  $G_2$  in fact preserves the special structure of  $l$ -factors: it maps the vertices of  $G_1$  onto consecutive vertices of  $G_2$ . That is very similar to the factor containment relation of words, which is the reason why we named these graphs  $l$ -factors.

**Lemma 5.3.5.** *Let  $G_1 = G_{w^{(1)},H_1}(1, 2, \dots, m_1)$  be a strong  $l$ -factor with letter partition  $\pi_1$ , and let  $G_2 = G_{w^{(2)},H_2}(1, 2, \dots, m_2)$  be an  $l$ -factor with letter partition  $\pi_2$ . If  $|p(\pi_1)| \geq |\pi_2|$ , then any embedding  $f : G_1 \rightarrow G_2$  satisfies either*

- (a)  $f(x) = f(1) - 1 + x$ ; or
- (b)  $f(x) = f(1) + 1 - x$ .

*Moreover, there is an embedding  $\eta : H_1 \rightarrow H_2$  such that*

- *in case (a),  $\eta(w^{(1)})$  appears as a factor of  $w^{(2)}$ ;*
- *in case (b),  $\eta(r(w^{(1)}))$  appears as a factor of  $w^{(2)}$ .*

*In particular, if  $H_1 = H_2 = H$ , then there is some  $\theta \in \text{Aut}(H) \times R$  such that  $\theta(w^{(1)})$  appears as a factor of  $w^{(2)}$ .*

*Proof.* By the assumptions of this lemma, Lemma 5.2.4 applies. We get that  $xy$  is an edge of  $\phi(G_1)$  if and only if  $f(x)f(y)$  is an edge of  $\phi(G_2)$ . Now, by the definition of  $l$ -factors the left hand side of the equivalence holds if and only if  $|x - y| = 1$ , and the right hand side holds if only if  $|f(x) - f(y)| = 1$ . As  $f$  is an injection we deduce that  $f(x) = f(1) + x - 1$  or  $f(x) = f(1) - x + 1$ . The rest follows directly from Lemma 5.2.4 as bags are injectively embedded into bags.  $\square$

Now we are ready to prove the last result of this section.

**Definition 5.3.4.** An infinite word  $w$  is called *ultimately almost periodic* if after removing some prefix of  $w$  the remaining word is almost periodic.

**Theorem 5.3.6.** Suppose  $\mathcal{P} = \mathcal{P}(w, H)$  for some infinite word  $w$  using all letters of a finite alphabet  $A$  and a prime graph  $H$  with  $V(H) = A$ . If  $\mathcal{P}$  is minimal above the Bell number, then  $w$  is ultimately almost periodic.

*Proof.* First of all, let  $A_f \subset A$  be the set of letters that appear finitely many times in  $w$  and let  $A' = A \setminus A_f$ . Take a finite prefix  $w_1 w_2 \dots w_k$  of  $w$  containing all the occurrences of the letters from  $A_f$  in  $w$ . Removing this prefix from  $w$  leaves us with the word  $w' = w_{k+1} w_{k+2} \dots$  on alphabet  $A'$  containing each letter of  $A'$  infinitely often. We prove that the subgraph  $H' = H[A']$  is prime. Suppose, for the sake of contradiction, that  $H'$  is not prime. Then there are two vertices  $a, b \in V(H') = A'$  such that each vertex of  $H'$  is adjacent either to both  $a$  and  $b$ , or to neither  $a$  nor  $b$ . However, as  $H$  is prime there is some vertex  $c \in V(H) \setminus V(H') = A_f$  which is adjacent to one of  $a, b$  but not the other. Now take an occurrence of  $c$ , say  $w_{k_c} = c$ , and a factor  $w_{k+1} w_{k+2} \dots w_{k'}$  of  $w$  in which each letter of  $A'$  appears at least  $5 \times 2^{l+1}$  times. We claim that the graph  $G = G_{w,H}(k_c, k+1, k+2, \dots, k')$ , which appears in  $\mathcal{P}(w, H)$ , does not appear in  $\mathcal{P}(w', H)$ . As  $\mathcal{P}(w', H)$  is above the Bell number and is contained in  $\mathcal{P}(w, H)$ , this will contradict the minimality of  $\mathcal{P}(w, H)$ . Let  $\pi$  be the letter partition of  $G' = G_{w,H}(k+1, k+2, \dots, k')$ , which is a strong  $(l_{\mathcal{P}}, 2)$  partition. Consider the partition  $p(\pi) = \{W_1, W_2, \dots, W_l\}$ . Notice that one of the bags, say,  $W_k$  contains both  $V_a$  and  $V_b$ . Also notice that every graph in  $\mathcal{P}(w', H)$  is a  $(|p(\pi)|, 2)$ -graph. In particular, if this class contains  $G$ , then  $G$  is a  $(|p(\pi)|, 2)$ -graph with some  $(|p(\pi)|, 2)$ -partition  $\pi_2$ . Applying Lemma 5.2.4 to the embedding of  $G'$  into  $G$ , we get that one bag  $W$  in  $\pi_2$  contains both  $V_a$  and  $V_b$ . But then vertex  $k_c \in V(G) \setminus V(G')$  has more than 2 neighbours and more than 2 non-neighbours in  $W$ . This contradicts the assumption that  $\pi_2$  is a  $(|p(\pi)|, 2)$ -partition and hence that  $G$  can be embedded into  $\mathcal{P}(w', H)$ .

Next we show that  $w'$  is almost periodic; that implies that  $w$  is ultimately almost periodic. For the sake of contradiction, suppose that  $w'$  is not almost periodic. Then there is a factor  $f$  and an infinite sequence of factors  $f_1, f_2, \dots$  of  $w'$  such that  $|f_i| < |f_{i+1}|$  and none of the  $f_i$ 's contains  $f$  as a factor. Let  $G_i = G_{f_i, H}(1, 2, \dots, |f_i|)$  for all  $i \in \mathbb{N}$ . By Lemma 5.3.4, we can find sequence of words  $g_1, g_2, \dots$  of increasing length, each of them a factor of some  $f_i$  such that the graphs  $G_i^s = G_{g_i, H}(1, 2, \dots, |g_i|)$  are strong  $l$ -factors for all  $i \in \mathbb{N}$ . Let  $\pi_i$  be the letter partition of  $G_i^s$ , which is a strong  $(l_{\mathcal{P}}, 2)$ -partition. Notice that for each  $i \in \mathbb{N}$  we have

$|p(\pi_i)| \in \{1, 2, \dots, l_P\}$ . Hence there exists  $k$  such that  $\mathcal{S}_k = \{G_i : |p(\pi_i)| = k\}$  is infinite. As  $\mathcal{P}$  is well-quasi-ordered, so is  $\mathcal{S}_k \subset \mathcal{P}$ . Hence  $\mathcal{S}_k$  contains an infinite chain  $G_{i_1} \subset G_{i_2} \subset \dots$ . From Lemma 5.3.5 we get that there is some  $\phi_k \in \text{Aut}(H') \times R$  such that  $\phi_k(g_{i_k})$  is a factor of  $(g_{i_{k+1}})$ . It is the same as saying that  $g_{i_k}$  is a factor of  $\phi_k^{-1}g_{i_{k+1}}$ , or denoting by  $\theta_1 = id$  the identity, and  $\theta_k = \phi_{k-1}^{-1}\phi_{k-2}^{-1}\dots\phi_1^{-1}$  for  $k \geq 2$ , we can say that  $\theta_k(g_{i_k})$  is a factor of  $\theta_{k+1}(g_{i_{k+1}})$ . Now, as the group  $\text{Aut}(H') \times R$  is finite, we get a subsequence  $\theta_{j_1}, \theta_{j_2}, \dots$  such that  $\theta_{j_i} = \theta$  for some fixed  $\theta \in \text{Aut}(H') \times R$ . Therefore, denoting  $g'_k = g_{i_{j_k}}$  we conclude that we have a subsequence  $g'_1, g'_2, \dots$  of  $g_1, g_2, \dots$  such that  $g'_i$  is a factor of  $g'_{i+1}$  for all  $i \in \mathbb{N}$ . Fix an embedding  $e_i$  from  $g'_i$  to  $g'_{i+1}$  and construct an infinite word  $w''$  as follows. Let  $w''_1 w''_2 \dots w''_{|g'_1|}$  be precisely the word  $g'_1$ . Let the next letters be formed by the suffix of  $g'_2$  following  $g'_1$  in the embedding  $e_1$  of  $g'_1$  into  $g'_2$ , then continue adding the suffix of  $g'_3$  following  $g'_2$  in the embedding  $e_2$  of  $w_2$  into  $g'_3$  and so on. If it happens that  $w''$  is not infinite, then form  $w''$  by reading the letters of  $g'_1$  from right to left and adding prefix of  $g'_2$  following  $g'_1$  in  $e_1$  (also reading from left to right), then prefix of  $g'_3$  and so on.

Finally, we show that the class  $\mathcal{P}(w'', H')$  is properly contained in  $\mathcal{P}(w', H')$ . Since  $\mathcal{P}(w'', H')$  is above the Bell number, this will finish the proof by contradicting the fact that  $\mathcal{P}(w', H')$  is minimal. In order to show the proper containment, we will construct a graph  $G \in \mathcal{P}(w', H') \setminus \mathcal{P}(w'', H')$ . Fix an embedding of  $f$  into  $w'$  and let  $f = w'_k w'_{k+1} \dots w'_{k'}$  and let  $n = \max(|g'_1|, k')$ . Let  $f' = w'_1 w'_2 \dots w'_n$ . Let  $S = \{\theta \in \text{Aut}(H') \times R : \theta(f')$  appears as a factor of  $w'\}$ . For all  $\theta \in S$ , let  $k_\theta$  be the least number such that  $\theta(f') = w'_{k_\theta+1} w'_{k_\theta+2} \dots w'_{k_\theta+n}$ . Now take  $n' = \max_{\theta \in S} \{k_\theta + n\}$  and let  $G = G_{w', H'}(1, 2, \dots, n')$ . Suppose  $G$  appears in  $\mathcal{P}(w'', H')$ . Then since  $G$  is a strong  $(l, 2)$ -graph and any graph in  $\mathcal{P}(w'', H')$  is an  $(l, 2)$ -graph we have that for each embedding  $G_{w'', H'}(i_1 i_2 \dots i_{n'})$  of  $G$  into  $w''$  we must have  $w''_{k+1} w''_{k+2} \dots w''_{k+n'} = \theta(w'_1 w'_2 \dots w'_n)$  for some  $\theta' \in \text{Aut}(H') \times R$ . Now, for any  $\theta \in S$  we have  $(\theta'\theta)(f') = \theta'(\theta(f')) = \theta'(w'_{k_\theta+1} w'_{k_\theta+2} \dots w'_{k_\theta+n})$  is a factor of  $w''_{k+1} w''_{k+2} \dots w''_{k+n'}$  and hence by construction  $w''$  it is a factor of  $w'$ . By the definition of  $S$  we see that  $\theta'\theta \in S$ . As  $S$  is finite we get that  $\{\theta'\theta : \theta \in S\} = S$ . Hence as  $S$  contains an identity, we get that  $\theta'\theta'' = id$  for some  $\theta'' \in S$ . We conclude that  $f'$  is a factor of  $w''$ . But as  $f$  is a factor of  $f'$  we conclude that  $f$  is a factor of  $w''$  which is clearly impossible by construction. Hence there is no embedding of  $G$  into  $\mathcal{P}(w'', H')$  and we are done.  $\square$

## 5.4 Decidability of the Bell number

As we mentioned in the introduction, every class below the Bell number can be characterised by a finite set of forbidden induced subgraphs. Therefore all classes for which the set of minimal forbidden induced subgraphs is infinite have speed above the Bell number. For classes defined by finitely many forbidden induced subgraphs, the problem of deciding whether their speed is above the Bell number is more complicated and decidability of this problem has been an open question. In this section, we employ our characterisation of minimal classes above the Bell number to answer this question positively.

Our main goal is to provide an algorithm that decides for an input consisting of a finite number of graphs  $F_1, \dots, F_n$  whether the speed of  $\mathcal{X} = \text{Free}(F_1, \dots, F_n)$  is above the Bell number. That is, we are interested in the following problem.

*Problem 5.4.1.*

INPUT: A finite set of graphs  $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$

OUTPUT: Yes, if the speed of  $\mathcal{X} = \text{Free}(\mathcal{F})$  is above the Bell number; no otherwise.

Our algorithm, following the characterisation of minimal classes above the Bell number, distinguishes two cases depending on whether the distinguishing number  $k_{\mathcal{X}}$  is finite or infinite. First we show how to discriminate between these two cases.

*Problem 5.4.2.*

INPUT: A finite set of graphs  $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$

OUTPUT: Yes, if  $k_{\mathcal{X}} = \infty$  for  $\mathcal{X} = \text{Free}(\mathcal{F})$ ; no otherwise.

**Theorem 5.4.1.** *There is a polynomial-time algorithm that solves Problem 5.4.2.*

*Proof.* By Theorem 5.3.1,  $k_{\mathcal{X}} = \infty$  if and only if  $\mathcal{X}$  contains one of the thirteen minimal classes listed there. By Theorem 5.3.3, each of the minimal classes is defined by finitely many forbidden induced subgraphs; thus membership can be tested in polynomial time. Then the answer to Problem 5.4.2 is no if and only if each of the minimal classes given by Theorem 5.3.1 contains at least one of the graphs in  $\mathcal{F}$ , which can also be tested in polynomial time.  $\square$

By Corollary 5.3.1, the minimal hereditary classes with finite distinguishing number with speed above the Bell number can be described as  $\mathcal{P}(w, H)$  with an almost periodic infinite word  $w$ . Here we give a more precise characterisation for classes defined by finitely many forbidden induced subgraphs.

**Definition 5.4.1.** Let  $w = w_1w_2\ldots$  be an infinite word over a finite alphabet  $A$ . If there exists some  $p$  such that  $w_i = w_{i+p}$  for all  $i \in \mathbb{N}$ , we call the word  $w$  *periodic* and the number  $p$  its *period*. If, moreover, for some period  $p$  the letters  $w_1, w_2, \ldots, w_p$  are all distinct, we call the word  $w$  *cyclic*.

If  $w$  is a finite word, then  $w' = (w)^\infty$  is the periodic word obtained by concatenating infinitely many copies of the word  $w$ ; thus  $w'_i = w_k$  for  $k = i \bmod |w'|$ .

A class  $\mathcal{X}$  of graphs is called a *periodic class* (*cyclic class*, respectively) if there exists a graph  $H$  with loops allowed and a periodic (cyclic, respectively) word  $w$  such that  $\mathcal{X} = \mathcal{P}(w, H)$ .

**Definition 5.4.2.** Let  $A = \{1, 2, \ldots, \ell\}$  be a finite alphabet,  $H$  a graph on  $A$  with loops allowed, and  $M$  a positive integer. Define a graph  $S_{H,M}$  with vertex set  $V(S_{H,M}) = A \times \{1, 2, \ldots, M\}$  and an edge between  $(a, j)$  and  $(b, k)$  if and only if one of the following holds:

- $ab \in E(H)$  and either  $|a - b| \neq 1$  or  $j \neq k$ ;
- $ab \notin E(H)$  and  $|a - b| = 1$  and  $j = k$ .

The graph  $S_{H,M}$  is called an  $(\ell, M)$ -*strip*.

Notice that a strip can be viewed as the graph obtained from the union of  $M$  disjoint paths  $(1, j) - (2, j) - \cdots - (\ell, j)$  for  $j \in \{1, 2, \ldots, M\}$  by swapping edges with non-edges between vertices  $(a, j)$  and  $(b, k)$  if  $ab \in E(H)$ .

**Theorem 5.4.2.** Let  $\mathcal{X} = \text{Free}(F_1, F_2, \ldots, F_n)$  with  $k_{\mathcal{X}}$  finite. Then the following conditions are equivalent:

- (a) The speed of  $\mathcal{X}$  is above the Bell number.
- (b)  $\mathcal{X}$  contains a periodic class.
- (c) For every  $p \in \mathbb{N}$ ,  $\mathcal{X}$  contains a cyclic class with period at least  $p$ .
- (d) There exists a cyclic word  $w$  and a graph  $H$  on the alphabet of  $w$  such that  $\mathcal{X}$  contains the  $\ell$ -factor  $G_{w,H}(1, 2, \ldots, 2\ell m)$  with  $\ell = |V(H)|$  and  $m = \max\{|F_i| : i \in \{1, 2, \ldots, n\}\}$ .
- (e) For any positive integers  $\ell, m$ , the class  $\mathcal{X}$  contains an  $(\ell, m)$ -strip.

*Proof.* (a)  $\Rightarrow$  (b): From Theorem 5.3.5 we know that  $\mathcal{X}$  contains a class  $\mathcal{P}(w, H)$  with some almost periodic word  $w = w_1w_2\ldots$  and a finite graph  $H$  with loops allowed. Let  $m = \max\{|F_1|, |F_2|, \ldots, |F_n|\}$  and let  $a = w_1w_2\ldots w_m$  be the word consisting of the first  $m$  letters of the infinite word  $w$ . Since  $w$  is almost periodic, the factor  $a$  appears in  $w$  infinitely often. In particular, there is  $m' > m$  such that  $w_{m'+1}w_{m'+2}\ldots w_{m'+m} = a$ . Define  $b$  to be the word between the two  $a$ 's in  $w$ , i.e., let  $b = w_{m+1}w_{m+2}\ldots w_{m'}$ . In this way,  $w$  starts with the initial segment  $aba$ .

We claim that  $\mathcal{X}$  contains the periodic class  $\mathcal{P}(w', H)$  with  $w' = (ab)^\infty$ . For the sake of contradiction, suppose that  $\mathcal{X}$  does not contain  $\mathcal{P}(w', H)$ . Then for some  $i \in \{1, 2, \dots, n\}$ , we have  $F_i \in \mathcal{P}(w', H)$ . So  $F_i \cong G_{w', H}(u_1, u_2, \dots, u_k)$  for some  $u_1 < u_2 < \dots < u_k$ . Let  $U = \{u_1, u_2, \dots, u_k\}$ . We are now looking for a monotonically increasing function  $f : U \rightarrow \mathbb{N}$  with these two properties: firstly,  $w_{f(u)} = w'_u$  for any  $u \in U$ ; secondly,  $f(u) - f(u') = 1$  if and only if  $u - u' = 1$ . If we can establish the existence of such a function, we will then have  $F_i \cong G_{w', H}(u_1, u_2, \dots, u_k) \cong G_{w, H}(f(u_1), f(u_2), \dots, f(u_k)) \in \mathcal{P}(w, H) \subseteq \mathcal{X}$ , a contradiction.

To construct such a function  $f$ , consider a maximal block  $\{u_j, u_{j+1}, \dots, u_{j+p}\}$  of consecutive integers in  $U$  (that is,  $u_{j-1} < u_j - 1$ ;  $u_j = u_{j+1} - 1$ ;  $\dots$ ;  $u_{j+p-1} = u_{j+p} - 1$ ;  $u_{j+p} < u_{j+p+1} - 1$ ). Furthermore, consider the word  $w'_{u_j} w'_{u_{j+1}} \dots w'_{u_{j+p}}$  of length at most  $m$ , which is a factor of  $w' = (ab)^\infty$  and thus also a factor of  $aba$  because  $|aba| > 2m$ . The word  $aba$ , being a factor of  $w$ , appears infinitely often in  $w$  because  $w$  is almost periodic. Hence not only we can define  $f : U \rightarrow \mathbb{N}$  in such a way that  $w_{f(u)} = w'_u$  for any  $u \in U$  and that blocks of consecutive integers in  $U$  are mapped to blocks of consecutive integers, but we can also do it monotonically and so that  $f(u) > f(u') + 1$  whenever  $u > u' + 1$ . This finishes the proof of the first implication.

(b)  $\Rightarrow$  (c): Let  $\mathcal{X}$  contain a class  $\mathcal{P}(w, H)$ , where  $w$  is a periodic word and  $H$  is a graph with loops allowed on the alphabet of  $w$ . If  $w$  is not cyclic (i.e., if some letters appear more than once within the period), the class  $\mathcal{P}(w, H)$  can be transformed into a cyclic one by extending the alphabet, renaming multiple appearances of the same letter within the period to different letters of the extended alphabet and modifying the graph  $H$  accordingly. More formally, let  $p$  be the period of  $w$ . Define a new word  $w' = (123 \dots p)^\infty$  and a graph  $H'$  with vertex set  $\{1, 2, \dots, p\}$  and an edge  $ij \in E(H')$  if and only if  $w_i w_j \in E(H)$ . Then any graph  $G_{w, H}(u_1, u_2, \dots, u_m) \in \mathcal{P}(w, H)$  is isomorphic to the graph  $G_{w', H'}(u_1, u_2, \dots, u_m) \in \mathcal{P}(w', H')$ . Hence  $\mathcal{X}$  also contains  $\mathcal{P}(w', H')$ , where the word  $w'$  is cyclic.

Since any periodic word of period  $p$  is also a periodic word of period  $kp$  for any  $k \in \mathbb{N}$ , by the same transformation a periodic class of period  $p$  can be transformed into a cyclic class of period  $kp$ .

(c)  $\Rightarrow$  (d): Follows directly from the definition of a cyclic class.

(d)  $\Rightarrow$  (a): Let  $\mathcal{X}$  contain the  $\ell$ -factor  $G_{w, H}(1, 2, \dots, 2\ell m)$ . We prove that  $\mathcal{X}$  then contains the whole class  $\mathcal{P}(w, H)$ ; its speed will then be above the Bell number by Proposition 5.3.1. For the sake of contradiction suppose that for some  $i$ , the graph  $F_i$  belongs to  $\mathcal{P}(w, H)$ . Then  $F_i \cong G_{w, H}(u_1, u_2, \dots, u_k)$  for some  $k \leq$



$m$  and  $u_1 < u_2 < \dots < u_k$ . Note that the period of  $w$  is  $\ell$  (by the definition of a cyclic class). Put  $u'_1 = u_1 \bmod \ell$ . Furthermore, for each  $i \geq 2$  let  $u'_i = (u_i \bmod \ell) + c_i \ell$ , where each  $c_i$  is chosen in such a way that  $0 < u'_i - u'_{i-1} \leq \ell + 2$  for all  $i$  and  $u'_i - u'_{i-1} = 1$  if and only if  $u_i - u_{i-1} = 1$ . By construction,  $F_i \cong G_{w,H}(u_1, u_2, \dots, u_k) \cong G_{w,H}(u'_1, u'_2, \dots, u'_k)$  and  $u'_k < k(\ell + 1) \leq m(\ell + 1) \leq 2\ell m$ . Hence  $F_i$  is isomorphic to an induced subgraph of  $G_{w,H}(1, 2, \dots, 2\ell m) \in \mathcal{X} = \text{Free}(F_1, F_2, \dots, F_n)$ , a contradiction.

(c)  $\Rightarrow$  (e): Let  $\mathcal{X}$  contain the cyclic class  $\mathcal{P}(w, H)$ , where  $w$  is a cyclic word with period  $p > \ell$ . Since  $p > \ell$ , the subgraph of  $G_{w,H}(1, 2, \dots, pM)$  induced by the bags corresponding to the first  $\ell$  letters of  $w$  is an  $(\ell, M)$ -strip.

(e)  $\Rightarrow$  (a): Let  $\ell_{\mathcal{X}}$ ,  $d_{\mathcal{X}}$  and  $c_{\mathcal{X}}$  be the constants given by Lemma 5.2.5 and put  $M = 2d_{\mathcal{X}} + c_{\mathcal{X}} + 2$ . We show that for any fixed positive integer  $\ell$ , the class  $\mathcal{X}$  contains a strong  $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph  $G$  such that its sparsification  $\phi(G)$  has a component of order at least  $\ell$ . Then we can apply Theorem 5.2.2.

So let  $\ell$  be a positive integer. By assumption,  $\mathcal{X}$  contains an  $(\ell, M)$ -strip  $S_{H,M}$ . By Lemma 5.2.5, after removing no more than  $c_{\mathcal{X}}$  vertices we are left with a strong  $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -graph  $S'$  with a strong  $(\ell_{\mathcal{X}}, d_{\mathcal{X}})$ -partition  $\pi$ . Let  $V_a = \{(a, j) \in V(S') : 1 \leq j \leq M\}$  be the letter bags of  $S'$ ,  $1 \leq a \leq \ell$ , and consider the prime partition  $p(\pi) = \{W_1, W_2, \dots, W_{\ell'}\}$ . If two vertices  $x, y$  belong to different bags of  $p(\pi)$ , then according to Lemma 5.2.3 we have  $|N(x) \ominus N(y)| \geq 5 \times 2^{\ell_{\mathcal{X}}} d_{\mathcal{X}} - 2d_{\mathcal{X}}$ . But notice that if we have two vertices  $(a, j), (a, j')$  of  $S'$  in the same letter bag  $V_a$ , then  $N(a, j) \ominus N(a, j') \subseteq \{(a-1, j), (a-1, j'), (a+1, j), (a+1, j')\}$ , so its size is at most 4. Hence, we deduce that  $V_a \subseteq W_{f(a)}$  for some function  $f$ .

Now notice that  $(V_a, V_b)$  is  $d_{\mathcal{X}}$ -dense, that is,  $ab$  is an edge of  $H$ , if and only if  $(W_{f(a)}, W_{f(b)})$  is  $d_{\mathcal{X}}$ -dense. Indeed, if one of them was  $d_{\mathcal{X}}$ -dense and the other  $d_{\mathcal{X}}$ -sparse, then  $(V_a, V_b)$  would be both  $d_{\mathcal{X}}$ -dense and  $d_{\mathcal{X}}$ -sparse, in which case  $|V_a| \leq 2d_{\mathcal{X}} + 1$ . But this is not true, as  $V_a$  is obtained from a set of size  $M = 2d_{\mathcal{X}} + 2 + c_{\mathcal{X}}$  by removing at most  $c_{\mathcal{X}}$  vertices.

It follows that  $\phi(S')$  is constructed by swapping edges with non-edges between  $V_a$  and  $V_b$  such that  $ab \in E(H)$ . Hence  $\phi(S')$  is a linear forest obtained from the paths  $(1, j) - (2, j) - \dots - (\ell, j)$  for  $j \in \{1, 2, \dots, M\}$  by removing at most  $c_{\mathcal{X}}$  vertices. As  $M > c_{\mathcal{X}}$ , at least one of the paths is left untouched. Therefore,  $\phi(S')$  contains a component of size at least  $\ell$ .  $\square$

Finally, we are ready to tackle the decidability of Problem 5.4.1.

### Algorithm 5.4.3.

INPUT: A finite set of graphs  $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$

OUTPUT: *Yes, if the speed of  $\mathcal{X} = \text{Free}(\mathcal{F})$  is above the Bell number; no otherwise.*

(1) *Using Theorem 5.4.1, decide whether  $k_{\mathcal{X}} = \infty$ . If it is, output yes and stop.*

(2) *Set  $m := \max\{|F_1|, |F_2|, \dots, |F_n|\}$  and  $\ell := 1$ .*

(3) *Loop:*

(3a) *For each graph (with loops allowed)  $H$  on  $\{1, 2, \dots, \ell\}$  construct the  $(\ell, \ell)$ -strip  $S_{H, \ell}$ . Check if some  $F_i$  is an induced subgraph of  $S_{H, \ell}$ . If for each  $H$  the strip  $S_{H, \ell}$  contains some  $F_i$ , output no and stop.*

(3b) *For each graph (with loops allowed)  $H$  on  $\{1, 2, \dots, \ell\}$  and for each word  $w$  consisting of  $\ell$  distinct letters from  $\{1, 2, \dots, \ell\}$  check if the  $\ell$ -factor  $G_{w^\infty, H}(1, 2, \dots, 2\ell m)$  contains some  $F_i$  as an induced subgraph. If one of these  $\ell$ -factors contains no  $F_i$ , output yes and stop.*

(3c) *Set  $\ell := \ell + 1$  and repeat.*

It remains to prove the correctness of this algorithm.

**Theorem 5.4.4.** *Algorithm 5.4.3 correctly solves Problem 5.4.1.*

*Proof.* We show that if the algorithm stops, it gives the correct answer, and furthermore that it will stop on any input without entering an infinite loop. First, if it stops in step (1), the answer is correct by [Balogh et al., 2005], since any class with infinite distinguishing number has speed above the Bell number.

Assume that the algorithm stops in step (3a) and outputs *no*. This is because every  $(\ell, \ell)$ -strip contains some forbidden subgraph  $F_i$ , hence no  $(\ell, \ell)$ -strip belongs to  $\mathcal{X}$ . By Theorem 5.4.2(e), the speed of  $\mathcal{X}$  is below the Bell number.

Next suppose that the algorithm stops in step (3b) and answers *yes*. Then  $\mathcal{X}$  contains the  $\ell$ -factor  $G_{w^\infty, H}(1, 2, \dots, 2\ell m)$ , where  $w^\infty$  is a cyclic word. Hence by Theorem 5.4.2(d) the speed of  $\mathcal{X}$  is above the Bell number.

Finally, if  $k_{\mathcal{X}} = \infty$  the algorithm stops in step (1). If  $k_{\mathcal{X}} < \infty$  and the speed of  $\mathcal{X}$  is above the Bell number, then by Theorem 5.4.2(d) the algorithm will stop in step (3b). If, on the other hand, the speed of  $\mathcal{X}$  is below the Bell number, then by Theorem 5.4.2(e) there exist positive integers  $\ell, M$  such that  $\mathcal{X}$  contains no  $(\ell, M)$ -strip. Let  $N = \max\{\ell, M\}$ . Obviously,  $\mathcal{X}$  contains no  $(N, N)$ -strip, because any  $(N, N)$ -strip contains some (many)  $(\ell, M)$ -strips as induced subgraphs and  $\mathcal{X}$  is hereditary. Therefore the algorithm will stop in step (3a) after finitely many steps.  $\square$

## 5.5 Deciding WQO for hereditary classes with finite distinguishing number

In this section we study well-quasi-ordering of hereditary classes  $\mathcal{X} = \text{Free}(\mathcal{F})$  with finite distinguishing number  $k_{\mathcal{X}}$ . In the first subsection we show that all classes below Bell number are:

- finitely defined
- well-quasi-ordered by labelled induced subgraph (and hence by induced subgraph) relation

In the second subsection we show the following. If a class  $\mathcal{X} = \text{Free}(\mathcal{F})$  with finite distinguishing number  $k_{\mathcal{X}}$  is above Bell number, then:

- $\mathcal{X}$  is not well-quasi-ordered by the labelled induced subgraph relation
- if  $\mathcal{F}$  is finite, then  $\mathcal{X}$  is not well-quasi-ordered by the induced subgraph relation
- if  $\mathcal{F}$  is infinite, then  $\mathcal{X}$  can be well-quasi-ordered by the induced subgraph relation

These results settle many of the questions relating well-quasi-ordering for the classes with finite distinguishing number. Indeed, the algorithm from the previous section could be used to detect whether a finitely defined class with finite distinguishing number is well-quasi-ordered by the induced subgraph relation. This is because the boundary of well-quasi-ordering for these classes coincides with the Bell number. Now, Allen et al. [2009] proved that all the classes below the Bell number are of bounded clique-width. As a consequence, for finitely defined classes with finite distinguishing number, well-quasi-ordering implies bounded clique-width, as conjectured by Daligault et al. [2010]. Finally, we remark that induced and labelled induced subgraph relations coincide for finitely defined classes with finite distinguishing number. Hence our conjecture from Chapter 3 holds for the family of classes with finite distinguishing number.

### 5.5.1 Classes below the Bell number

In this subsection we will prove that each class below the Bell number is well-quasi-ordered by showing that each of them is a subclass of  $k$ -uniform graphs for some value of  $k$ . The classes of  $k$ -uniform graphs were first introduced in [Korpelainen and Lozin, 2011a] and the definition is as follows.

Let  $k$  be a natural number,  $K$  a symmetric 0-1 square matrix of order  $k$  and  $F_k$  be a graph on  $\{1, 2, \dots, k\}$ . Let  $G$  be the disjoint union of infinitely many copies of  $F_k$  and for  $i = 1, 2, \dots, k$  let  $V_i$  be a subset of  $V(H)$  containing vertex  $i$  from each copy of  $F_k$ . Construct from  $H$  an infinite graph  $H(K)$  on the same vertex set by connecting two vertices  $u \in V_i$  and  $v \in V_j$  if and only if  $uv \in E(H)$  and  $K(i, j) = 0$  or  $uv \notin E(H)$  and  $K(i, j) = 1$ . Finally, let  $\mathcal{P}(K, F_k)$  be the hereditary class consisting of all the finite induced subgraphs of  $H(K)$ .

**Definition 5.5.1.** A graph  $G$  is called  $k$ -uniform if  $G \in \mathcal{P}(K, F_k)$  for some  $K$  and  $F_k$ .

The same paper proves the following two facts about  $k$ -uniform graphs:

**Theorem 5.5.1.** *If a graph  $G$  has a subset  $W$  of at most  $c$  vertices such that  $G - W$  is  $k$ -uniform, then  $G$  is  $2^c(k + 1) - 1$ -uniform.*

**Theorem 5.5.2.** *For any fixed  $k$ , the set of  $k$ -uniform graphs is well-quasi-ordered by the labelled-induced subgraph relation.*

To prove the main result of this section we need the following technical lemma.

**Lemma 5.5.1.** *Let  $G$  be a strong  $(l, d)$ -graph with sparsification  $\phi(G)$  containing components of size at most  $m$ . Then  $G$  is a  $k$ -uniform graph for  $k \leq ml^m 2^{\binom{m}{2}+1}$ .*

*Proof.* Let  $\pi = \{V_1, V_2, \dots, V_{l^m}\}$  be a strong  $(l, d)$ -partition of  $G$ . Two components  $C$  and  $C'$  of  $\phi(G)$  will be called  $\pi$ -equivalent if there is an isomorphism  $f : C \rightarrow C'$  such that for every vertex  $v \in C$  both  $v$  and  $f(v)$  belong to the same bag  $V_i$  of the partition  $\pi$ . Let  $V'$  be a subset of vertices of  $V(G)$  containing exactly one connected component from each  $\pi$ -equivalence class. Notice that each graph on  $m$  vertices can be embedded into bags of partition  $\pi$  in at most  $l^m \leq l^m$  different (not  $\pi$ -equivalent) ways. As there are at most  $2^{\binom{m}{2}+1}$  graphs on at most  $m$  vertices we obtain  $|V'| \leq ml^m 2^{\binom{m}{2}+1}$ . Let  $k = |V'|$  and let  $F_k = \phi(G)[V']$ . For two vertices  $u, v \in V'$ ,  $u \in V_i$ ,  $v \in V_j$  define  $K(u, v) = 0$  if  $(V_i, V_j)$  is  $d$ -sparse and  $K(u, v) = 1$  if  $(V_i, V_j)$  is  $d$ -dense. It is not hard to see that  $\mathcal{P}(K, F_k)$  contains  $G$ , so  $G$  is  $k$ -uniform with  $k$  as stated and we are done.  $\square$

**Theorem 5.5.3.** *Every hereditary class below Bell number is well-quasi-ordered by the labelled induced subgraph relation.*

*Proof.* Let  $\mathcal{X}$  be a hereditary class below Bell number. Then  $\mathcal{X}$  must have finite distinguishing number  $k_{\mathcal{X}}$ . Let  $G \in \mathcal{X}$ . By Theorem 5.2.5 we know that removal of

at most  $c_{\mathcal{X}}$  vertices leaves us with a strong  $(l_X, d_X)$  graph  $G'$ . By Theorem 5.2.2 it follows that there is an absolute constant  $m_{\mathcal{X}}$  dependent on  $\mathcal{X}$  only such that the sparsification of  $G'$  has a component of size at most  $m_{\mathcal{X}}$ . From Lemma 5.5.1 it follows that  $G'$  is  $k'$ -uniform graph with  $k' = m_{\mathcal{X}} l_{\mathcal{X}}^{m_{\mathcal{X}}} 2^{\binom{m_{\mathcal{X}}}{2}+1}$ . By Theorem 5.5.1  $G$  is  $k = 2^{c_X}(k' + 1) - 1$ -uniform. Hence there is a fixed constant  $k$ , such that every graph of the class  $X$  is  $k$ -uniform. Hence by  $\mathcal{X}$  is a subclass of  $k$ -uniform graphs and by Theorem 5.5.2 we deduce that  $X$  is well-quasi-ordered.  $\square$

In [Daligault et al., 2010] it was proved that all labelled well-quasi-ordered hereditary graph classes are finitely defined. Together with Theorem 5.5.3 we obtain the following interesting fact.

**Corollary 5.5.1.** *Every hereditary class below Bell number is finitely defined.*

It follows that given any hereditary class, we can decide whether it is above Bell number or below. If the class  $X$  is infinitely defined then by the Corollary 5.5.1 it follows that it is above Bell number. Otherwise, we apply the Algorithm 5.4.3 from the previous section.

### 5.5.2 Classes above the Bell number with finite distinguishing number

The following three theorems explore well-quasi-ordering of the classes above the Bell number with finite distinguishing number.

**Theorem 5.5.4.** *Suppose a class  $\mathcal{X}$  with finite distinguishing number  $k_{\mathcal{X}}$  is above the Bell number. Then  $\mathcal{X}$  is not well-quasi-ordered by labelled induced subgraph relation.*

*Proof.* From Theorem 5.3.5 we obtain that  $\mathcal{X}$  contains  $P(w, H)$  with  $w = a_1 a_2 \dots$  almost periodic and  $H$  prime. Let  $M$  be the smallest number such that each letter of  $w$  appears at least  $5 \times 2^{|H|}$  times. It is possible to pick such  $M$  because  $w$  is almost periodic. For  $i \geq M$ , let  $G_i = G_{w, H}(1, 2, \dots, i)$ . Then  $G_i$  is a strong  $(2, |H|)$ -graph with its letter partition and  $\phi(G_i) = P_i$  - a path with  $i$  vertices. By Lemma 5.2.4 all the embeddings  $G_i \rightarrow G_j$  must correspond to embeddings  $P_i \rightarrow P_j$  and must respect the labels of the vertices. However, the paths form an antichain with respect to labelled induced subgraph relation. This implies that graphs  $G_i$  form an antichain with respect to labelled induced subgraph relation as well. So  $\mathcal{X}$  is not well-quasi-ordered by the labelled induced subgraph relation.  $\square$

**Theorem 5.5.5.** *Suppose a class  $\mathcal{X} = \text{Free}(F_1, F_2, \dots, F_n)$  is above Bell number and has finite distinguishing number. Then  $\mathcal{X}$  is not well-quasi-ordered by the induced subgraph relation.*

*Proof.* Let  $\mathcal{X}$  be as in the statement. By Lemma 5.4.2 (c) we obtain that  $\mathcal{X}$  contains a periodic class  $P(w, H)$  with some  $w = (a_1 a_2 \dots a_k)^\infty$  and we can assume that  $H$  is prime. Let  $m = \max(|F_1|, |F_2|, \dots, |F_n|)$  and let  $M = \max(5 \times 2^{k+1}, m + 1)$ . For each  $i \geq M$ , let  $G_i$  be a graph obtained from  $G_{w,H}(1, 2, \dots, ik)$  by swapping edge with non-edge, between vertices 1 and  $ik$ . More formally, let  $G_i$  be a graph on vertex set  $[ik]$  with vertices  $j < j'$  adjacent to each other if and only if  $(j, j') \neq (1, k)$  and  $jj' \in E(G_{w,H}(1, 2, \dots, ik))$  or  $(j, j') = (1, ik)$  and  $jj' \notin E(G_{w,H}(1, 2, \dots, ik))$ .

Notice that each  $G_i$  is a strong  $(|H|, 2)$  graph, with  $|H|$  bags in the prime partition, and its sparsification  $\phi(G_i)$  is a cycle  $C_{ik}$ . Therefore, by Lemma 5.2.4 the set of graphs  $S = \{G_i : i \geq M\}$  is an antichain.

We finish the proof by showing that  $S \subset \mathcal{X}$ . Notice that deleting a vertex  $l$ , for some  $1 \leq l \leq ik$  of  $G_i$  we are left with a graph isomorphic to  $G_{w,H}(l + 1, l + 2, \dots, l + ik - 1)$  which is in  $P(w, H)$  and hence in  $\mathcal{X}$ . We deduce that all proper induced subgraphs of  $G_i$  are in  $\mathcal{X}$ . Now for all  $j$ ,  $H_j \notin \mathcal{X}$  and  $|G_i| > |H_j|$ . Hence  $H_j$  is not an induced subgraph of  $G_i$ . As this holds for all  $j$ , we have  $G_i \in \text{Free}(F_1, F_2, \dots, F_n)$ . So for all  $i \geq M$ ,  $G_i \in \mathcal{X}$ , hence  $S \subset \mathcal{X}$ .  $\square$

**Theorem 5.5.6.** *Each minimal class above the Bell number  $\mathcal{X} = \mathcal{P}(w, H)$  with  $w$  almost periodic and  $H$  prime is well-quasi-ordered by the induced subgraph relation.*

*Proof.* Suppose we have an antichain  $S \in \mathcal{X}$ . Then for any  $G \in S$  we have that  $S \setminus \{G\}$  is an antichain in  $\text{Free}(G) \cap \mathcal{X}$ . The class  $\text{Free}(G) \cap \mathcal{X}$  is below Bell number and hence by Theorem 5.5.3, it is well-quasi-ordered. Hence  $S \setminus \{G\}$  is finite which means that  $S$  is finite. Hence  $\mathcal{X}$  is well-quasi-ordered by the induced subgraph relation.  $\square$

We end this section with conjecture on which hereditary classes (not necessarily finitely defined) with finite distinguishing number are well-quasi-ordered by the induced subgraph relation. The conjecture is based on the results we obtained in this section and the result of Guoli Ding for monotone classes (closed under taking subgraphs). In [Ding, 1992] the author proved that a monotone class is well-quasi-ordered if and only if it contains finitely many graphs from two antichains: the set of cycles and the set of so-called  $H$ -graphs. Let us call the antichains produced in the proof of Theorem 5.5.5 the antichains of generalised cycles. Define the antichains of generalised  $H$ -graphs in a similar way.

**Conjecture 5.5.1.** A hereditary class with finite distinguishing number is well-quasi-ordered by the induced subgraph relation if and only if it contains only finitely many elements from every antichain of generalised cycles and generalised  $H$ -graphs.

## 5.6 Conclusion

In this part, we have characterised all minimal hereditary classes of graphs whose speed is at least the Bell number  $B_n$ . This characterisation allowed us to show that the problem of determining if the speed of a hereditary class  $\mathcal{X}$  defined by finitely many forbidden induced subgraphs is above or below the Bell number is decidable, i.e., there is an algorithm that gives a solution to this problem in a finite number of steps. However, the complexity of this algorithm, in terms of the input forbidden graphs, remains an open question. In particular, it would be interesting to determine if there is a polynomial bound on the minimum  $\ell$  such that the input class  $\mathcal{X}$  contains an  $\ell$ -factor as in Theorem 5.4.2(d) if it is above the Bell number, and it fails to contain any  $(\ell, \ell)$ -strip as in Theorem 5.4.2(e) if it is below.

We also verified the conjecture of Daligault et al. [2010] and proved decidability of well-quasi-ordering for classes of finite distinguishing number. The boundary of well-quasi-ordering coincided with the boundary separating classes of  $k$ -uniform graphs from the classes for which the uniformicity is unbounded.

To settle the conjecture of Daligault et al. [2010] and to obtain results about deciding well-quasi-ordering for classes with infinite distinguishing number, we suggest to study letter graphs which are known to be well-quasi-ordered and try to identify the minimal classes of unbounded lettericity. The classes of unbounded lettericity should contain infinite antichains similar to the ones we obtained in this Chapter and the ones we described in Section 2.6. A further indication that this is a promising direction is due to a certain orthogonality between the notions of uniformicity and lettericity. Indeed, the building blocks of  $k$ -uniform graphs consist of bags which are independent set of cliques with possible matchings or comatchings between them, but a class of matchings has unbounded lettericity. On the other hand, the letter graphs are obtained when instead of matchings and comatchings we use chain graphs, but a class of chain graphs has unbounded uniformicity. Finally, we note that despite the similarity to the question we solved in this section, the question for letter graphs is still a very challenging research question. In particular, adapting our approach to letter graphs, one would need to develop sparsification tools for chain graphs which seems to be a hard task.

## Chapter 6

# Subquadratic properties

We say that a hereditary graph property  $X$  is *subquadratic* if there is a function  $f(n) = o(n^2)$  bounding the number of edges in all  $n$ -vertex graphs in  $X$ . The family of subquadratic properties contains many important classes such as graphs of bounded vertex degree, of bounded tree-width, all proper minor closed graph classes. In all these examples, the number of edges is bounded by a linear function in the number of vertices and all of the listed properties are rather small (see e.g. [Norine et al., 2006] for the number of graphs in proper minor closed graph classes). In the terminology of [Balogh et al., 2000], they all are at most factorial. The family of subquadratic properties is much wider and contains classes with a superfactorial speed of growth, such as projective plane graphs (or more generally  $C_4$ -free bipartite graphs), in which case the number of edges is  $\Theta(n^{\frac{3}{2}})$ . In fact, as we show in Section 6.1, subquadratic properties have a nice structural characterization: these are precisely hereditary classes of graphs without large bicliques as subgraphs.

Our main result of this Chapter is proved in Section 6.2: a subquadratic property which is well-quasi-ordered by the induced subgraph relation is of bounded tree-width. Since bounded tree-width implies bounded clique-width, this verifies (and even strengthens) the conjecture of Daligault et al. [2010] for subquadratic properties. In Section 6.3 we further show that for a finitely defined subquadratic property, it is a decidable task to test well-quasi-orderability by the induced subgraphs relation. As a byproduct, in Section 6.4 we show that computing a biclique of order  $k$  in a  $P_s$ -free graphs is fixed parameter linear when parameterised by  $k$  and  $s$ .

All preliminary information related to the topic of the chapter can be found in Section 6.1, where we also prove a number of preparatory results.



## 6.1 Notations and preparatory results

### 6.1.1 Infinite antichains

Recall the graph  $H_i$  represented in Figure 1.1). By connecting two vertices of degree one having a common neighbour in  $H_i$ , we obtain a graph represented on the left of Figure 6.1. Let us denote this graph by  $H'_i$ . By further connecting the other pair of vertices of degree one we obtain the graph  $H''_i$  represented on the right of Figure 6.1.

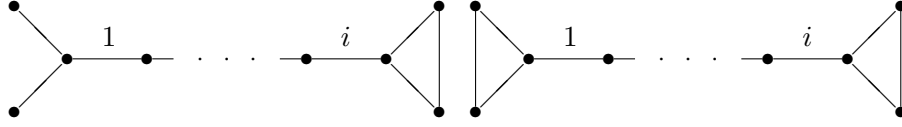


Figure 6.1: Graphs  $H'_i$  and  $H''_i$

We call any graph of the form  $H_i$ ,  $H'_i$  or  $H''_i$  an  $H$ -graph. Furthermore, we will refer to  $H''_i$  a *tight*  $H$ -graph and to  $H'_i$  a *semi-tight*  $H$ -graph. In an  $H$ -graph, the path connecting two vertices of degree 3 will be called the *body* of the graph, and the vertices which are not in the body the *wings*.

Following standard graph theory terminology, we call a chordless cycle of length at least four a *hole*. Let us denote by

$\mathcal{C}$  the set of all holes and all  $H$ -graphs.

It is not difficult to see that any two distinct (i.e. non-isomorphic) graphs in  $\mathcal{C}$  are incomparable with respect to the induced subgraph relation. In other words,

**Claim 6.1.1.**  $\mathcal{C}$  is an antichain with respect to the induced subgraph relation.

Moreover, from the poof of Theorem 6.2.1 we will see that for subquadratic classes of unbounded tree-width this antichain is unavoidable, or *canonical*, in the terminology of [Ding, 2009]. Suggested by this observation, we introduce the following definition.

**Definition 6.1.1.** The graphs in the set  $\mathcal{C}$  will be called **CANONICAL**.

The *order* of a canonical graph  $G$  is either the number of its vertices, if  $G$  is a hole, or the number of vertices in its body, if  $G$  is an  $H$ -graph.

### 6.1.2 Excluded graphs in subquadratic classes

In the current and the following subsections we obtain several important Ramsey-type results concerning subquadratic properties. By  $R = R(k, r, m)$ , we denote the Ramsey number, i.e. the minimum  $R$  such that in every coloring of  $k$ -subsets of an  $R$ -set with  $r$  colors there is a monochromatic  $m$ -set, i.e. a set of  $m$  elements all of whose  $k$ -subsets have the same color.

It is obvious that any subquadratic class must exclude at least one complete graph and at least one complete bipartite graph, since the number of edges in these graphs is proportional to the square of the number of vertices. It turns out that this restriction is not only necessary but also sufficient for a hereditary class to be subquadratic.

**Theorem 6.1.1.** *Let  $X$  be a hereditary class. Then the following three statements are equivalent.*

- (1)  $X$  is subquadratic,
- (2) there exist numbers  $t$  and  $q$  such that  $K_t \notin X$  and  $K_{q,q} \notin X$ ,
- (3) there exists a number  $p$  such that no graph in  $X$  contains  $K_{p,p}$  as a subgraph (not necessarily induced).

*Proof.* The implication (1)→(2) is obvious, since classes containing all cliques or all bicliques are not subquadratic. The implication (2)→(3) follows with  $p$  equal  $R(2, 2, \max(t, q))$ , where  $R$  is the Ramsey number. Now let us show that (3) implies (1).

Let  $Z_{r,s}(n_1, n_2)$  be the Zarankiewicz number, i.e. the smallest integer  $k$  such that every bipartite graph that has  $n_1$  vertices on one side of its bipartition,  $n_2$  vertices on the other side, and  $k$  edges contains a subgraph isomorphic to  $K_{r,s}$ . Kővári et al. [1954] showed that

$$Z_{r,s}(n_1, n_2) < (r-1)^{1/s}(n_1-s+1)n_2^{1-1/s} + (s-1)n_2.$$

From this we conclude that the number of edges in  $n$ -vertex graphs in  $X$  is less than  $4pn^{2-1/p}$ . Indeed, assume to the contrary that an  $n$ -vertex graph  $G \in X$  contains at least  $4pn^{2-1/p}$  edges. Then it must contain a bipartite subgraph  $G'$  (not necessarily induced) with at least  $2pn^{2-1/p}$  edges. Let  $n_1$  and  $n_2$  be the number of vertices in the two parts of  $G'$ . Since  $n_1 < n$ ,  $n_2 < n$  and  $(p-1)^{1/p} < 2$ , we conclude with the help of the above bound on the Zarankiewicz number that

$$2pn^{2-1/p} > 2n^{2-1/p} + (p-1)n > (p-1)^{1/p}(n_1-p+1)n_2^{1-1/p} + (p-1)n_2 > Z_{p,p}(n_1, n_2).$$

By the definition of the Zarankiewicz number, this implies that  $G'$ , and hence  $G$ , contains a subgraph isomorphic to  $K_{p,p}$ , contradicting (3). Therefore, (3) implies that the number of edges in  $n$ -vertex graphs in  $X$  is less than  $4pn^{2-1/p}$ , i.e.  $X$  is subquadratic.  $\square$

### 6.1.3 Long paths in graphs in subquadratic classes

In this subsection, we prove that graphs in a subquadratic class containing a large path also contain a large induced (i.e. chordless) path. We start with the following auxiliary result, where by  $P(r, m)$ , we denote the minimum  $n$  such that in every coloring of the elements of an  $n$ -set with  $r$  colors there exists a subset of  $m$  elements of the same color (the pigeonhole principle).

**Lemma 6.1.1.** *For each  $p$  and  $q$  there is a number  $C = C(p, q)$  such that whenever a graph  $G$  contains two families of sets  $\mathcal{A} = \{V_1, V_2, \dots, V_C\}$  and  $\mathcal{B} = \{W_1, W_2, \dots, W_C\}$  with all sets being disjoint of size  $p$  and with at least one edge between every two sets  $V_i \in \mathcal{A}$  and  $W_j \in \mathcal{B}$ , then  $G$  contains a biclique  $K_{q,q}$ .*

*Proof.* We define  $r := P(p^q, q)$  and  $C(p, q) := P(p^r, q)$  and consider an arbitrary collection  $A$  of  $r$  sets from  $\mathcal{A}$ . Since each set in  $\mathcal{B}$  has a neighbor in each set in  $\mathcal{A}$ , the family of the sets in  $\mathcal{B}$  can be colored with at most  $p^r$  colors so that all sets of the same color have a common neighbor in each of the  $r$  chosen sets of collection  $A$ . By the choice of  $C(p, q)$ , one of the color classes contains a collection  $B$  of at least  $q$  sets. For each set in  $A$ , we choose a vertex which is a common neighbor for all sets in  $B$  and denote the set of  $r$  chosen vertices by  $U$ . The vertices of  $U$  can be colored with at most  $p^q$  colors so that all vertices of the same color have a common neighbor in each of the  $q$  sets of collection  $B$ . By the choice of  $r$ ,  $U$  contains a color class  $U_1$  of least  $q$  vertices. For each set in  $B$ , we choose a vertex which is a common neighbor for all vertices of  $U_1$  and denote the set of  $q$  chosen vertices by  $U_2$ . Then  $U_1$  and  $U_2$  form a biclique  $K_{q,q}$ .  $\square$

**Theorem 6.1.2.** *For every  $s$  and  $q$  there is a number  $Y = Y(s, q)$  such that every graph with a path of length at least  $Y$  contains either a path  $P_s$  as an induced subgraph or a biclique  $K_{\lfloor q/2 \rfloor, \lfloor q/2 \rfloor}$  as a (not necessarily induced) subgraph.*

*Proof.* We use induction on  $s$  and  $q$ . For  $s = 1$  and arbitrary  $q$  or for  $q = 1$  and arbitrary  $s$ , we can take  $Y(s, q) = 1$ . So assume  $s > 1$  and  $q > 1$ . Let  $t = Y(s, q - 1)$  and  $k = Y(s - 1, C(t, q))$ . Both numbers must exist by the induction hypothesis.

Consider a graph  $G$  with a path  $P = v_1 v_2 \dots v_{kt}$  on  $kt$  vertices and split  $P$  into  $k$  subpaths on  $t$  vertices each. We denote the vertices of the  $i$ -th subpath by

$V_i$  and form a graph  $H$  on  $k$  vertices  $\{h_1, h_2, \dots, h_k\}$  in which  $h_i h_j$  is an edge if and only if there is an edge in  $G$  joining a vertex of  $V_i$  to a vertex of  $V_j$ . Since  $h_i$  is joined to  $h_{i+1}$  for each  $i = 1, \dots, k-1$ , the graph  $H$  has a path on  $k$  vertices, and since  $k = Y(s-1, C(t, q))$ , it has either an induced path on  $s-1$  vertices or a biclique of order  $C(t, q)$ . In the graph  $G$ , the latter case corresponds to two families of  $C(t, q)$  pairwise disjoint subsets with  $t$  vertices in each subset and with an edge between any two subsets from different families. Therefore, Lemma 6.1.1 applies proving that  $G$  contains a biclique  $K_{q,q}$ .

Now assume  $H$  contains an induced path  $P_{s-1}$ . In the graph  $G$ , this path corresponds to an ordered sequence of subsets  $V_{i_1}, V_{i_2}, \dots, V_{i_{s-1}}$  with edges appearing only between consecutive subsets of the sequence. Therefore, in the subgraph of  $G$  induced by these subsets, any vertex  $v$  in  $V_{i_1}$  is of distance at least  $s-2$  from any vertex  $u$  in  $V_{i_{s-1}}$ . If the distance between  $v$  and  $u$  is  $s-1$ , the graph  $G$  has an induced path  $P_s$  and we are done. So, assume the distance between any two vertices of  $V_{i_1}$  and  $V_{i_{s-1}}$  is exactly  $s-2$ , and consider a path with exactly one vertex  $w_p$  in each  $V_{i_p}$ .

If vertex  $w_1$  has a neighbour  $w \in V_{i_1}$  which is not adjacent to  $w_2$ , then  $ww_1w_2 \dots w_{s-1}$  is an induced path  $P_s$  and we are done. Therefore, we must assume that  $w_2$  is adjacent to every vertex of  $V_{i_1}$ , since this set induces a connected subgraph. As the size of  $V_{i_1}$  is  $t = Y(s, q-1)$ , it contains either an induced path  $P_s$ , in which case we are done, or a biclique  $K_{\lfloor (q-1)/2 \rfloor, \lceil (q-1)/2 \rceil}$ . In the latter case, the biclique together with  $w_2$  form a biclique of the desired size  $K_{\lfloor q/2 \rfloor, \lceil q/2 \rceil}$ , so we are done as well. This completes the proof.  $\square$

Taking into account that a large biclique gives rise either to a large induced biclique or a large clique, Theorem 6.1.2 can also be restated as follows.

**Theorem 6.1.3.** *For every  $s, t$ , and  $q$ , there is a number  $Z = Z(s, t, q)$  such that every graph with a path of length at least  $Z$  contains either  $P_s$  or  $K_t$  or  $K_{q,q}$  as an induced subgraph.*

#### 6.1.4 Some bounds on the number $Z(s, t, q)$

Theorem 6.1.3 proves the existence of the number  $Z(s, t, q)$  but does not tell anything about its value. Deriving any reasonable bounds on  $Z(s, t, q)$  seems to be a non-trivial problem. To make a progress in this direction, in this section we restrict ourselves to the case  $t = 3, q = 2$ . These are the minimal values of  $t$  and  $q$  for which the problem is still non-trivial. On the other hand, this restriction allows us to derive the following lower bound on  $Z(s, t, q)$ .

**Theorem 6.1.4.** *For sufficiently large values of  $s$ , we have  $Z(s, 3, 2) > s^2/64$ .*

*Proof.* To prove the theorem, we will show that  $Z(8n + 1, 3, 2) > n^2 + 1$  for  $n \geq 6$ . To this end, we will construct a  $(K_3, K_{2,2})$ -free graph  $G$  containing a path with  $n^2 + 1$  vertices and containing no induced path on  $8n + 1$  vertices. To simplify the description of  $G$  we will additionally assume that  $n$  is even.

We start with a path  $P = (0, 1, 2, \dots, n^2)$  with  $n^2 + 1$  vertices listed along the path. Let us call the vertices of  $P$  of the form  $kn$  with  $k \in \{2, 4, \dots, n - 2\}$  *special*. In addition to the edges of the path, we introduce a number of *chordal edges* (or simply *chords*) by connecting each special vertex  $kn$  to the following odd vertices:  $k + 3, 2n + k + 3, 4n + k + 3, \dots, n^2 - 2n + k + 3$ . Clearly, the graph constructed in this way is bipartite, and hence is  $K_3$ -free. Let us show that it is  $K_{2,2}$ -free. To contain a  $K_{2,2}$ , the graph must contain two even vertices whose neighbourhoods share at least two odd vertices. Clearly, any two non-special even vertices share at most one neighbour. Also, it is not difficult to see that for  $n \geq 6$  any special even vertex has at most one common neighbour with any other even vertex of the graph. Hence,  $G$  is  $K_{2,2}$ -free.

To finish the proof, consider an arbitrary induced path  $P^*$  in  $G$  and assume it contains five odd vertices with the same residue modulo  $2n$ . We denote these vertices by  $a_i = 2np_i + a$ ,  $i = 1, 2, \dots, 5$ ,  $0 < a < 2n$ , in the order of their appearances in  $P^*$ . Since the distance between  $a_i$  and  $a_{i+1}$  in  $P$  is at least  $2n$ , the subpath of  $P$  connecting  $a_i$  to  $a_{i+1}$  contains at least one chord, and hence  $P^*$  contains at least one chordal edge between  $a_i$  and  $a_{i+1}$ . Let us consider the set  $A$  of all odd-numbered endpoints of the chordal edges in  $P^*$  and let us denote by  $B = \{b_1, b_2, \dots, b_k\}$  the set of all distinct residues modulo  $2n$  of the vertices in  $A$  (we assume that the elements of  $B$  are listed in the increasing order). Then

- either  $a \in B$ , in which case all five chordal edges incident to the vertices  $a_1, \dots, a_5$  must be present in  $P^*$ , which is impossible since otherwise  $P^*$  has a vertex of degree at least 5,
- or  $b_i < a < b_{i+1}$  for some  $i = 1, \dots, k - 1$ . Consider the subpath  $P'$  of  $P$  between the vertices  $2np_2 + b_i$  and  $2np_2 + b_{i+1}$ . This subpath contains vertex  $a_2$  of  $P^*$ . Therefore, the entire subpath  $P'$  must belong to  $P^*$ . To see this, observe that no even vertex of this subpath is special and hence leaving or entering this subpath through an even vertex is impossible. Leaving or entering this subpath through an odd vertex different from  $b_i$  and  $b_{i+1}$  is also impossible as well. Indeed, leaving or entering  $P'$  through an odd vertex different from  $b_i$  and  $b_{i+1}$  could only be possible by means of a chordal edge,

in which case  $B$  would contain an element between  $b_i$  and  $b_{i+1}$ . Therefore,  $P'$  entirely belongs to  $P^*$ . Similarly, the following two subpaths of  $P$  entirely belong to  $P^*$ :  $(2np_3 + b_i, 2np_3 + b_{i+1})$  and  $(2np_4 + b_i, 2np_4 + b_{i+1})$ . Finally, since  $b_i$  belongs to  $B$ , the unique special even vertex incident to  $2np_2 + b_i$ ,  $2np_3 + b_i$ ,  $2np_4 + b_i$  also belongs to  $P^*$ . But then  $P^*$  contains a vertex of degree at least 3 which is impossible.

- or  $a < b_1$  or  $b_k < a$ . If  $b_k < a$ , then by analogy with the previous case we conclude that  $P^*$  entirely contains the following three subpaths of  $P$ :  $(2np_2 + b_k, 2np_2 + 2n)$ ,  $(2np_3 + b_k, 2np_3 + 2n)$ , and  $(2np_4 + b_k, 2np_3 + 2n)$ . This is impossible, since otherwise  $P^*$  contains a vertex of degree at least 3. Similarly, the case  $a < b_1$  is impossible.

Since the above cases exhaust all possibilities for  $a$  and  $B$ , we conclude that  $P^*$  contains at most 4 odd vertices with the same residue modulo  $2n$ . Therefore, the number of odd vertices in  $P^*$  is at most  $4n$ , and hence the total number of vertices in  $P^*$  is at most  $8n + 1$ .

We have proved that  $Z(8n + 1, 3, 2) > n^2 + 1$  for  $n \geq 6$ . Therefore, for sufficiently large values of  $s$ , we have  $Z(s, 3, 2) > s^2/64$ .  $\square$

## 6.2 Main result: WQO classes of bounded treewidth

As we mentioned in the introduction, Daligault et al. [2010] conjectured that every hereditary class which is well-quasi-ordered by the induced subgraph relation is of bounded clique-width. In the present section, we verify this conjecture for subquadratic classes. Moreover, we prove an even stronger result stating that every subquadratic class which is well-quasi-ordered by the induced subgraph relation is of bounded tree-width. Together with the well-known fact that the tree-width of cliques and bicliques is not bounded, we obtain the following dichotomy result.

**Theorem 6.2.1.** *A hereditary class which is well-quasi-ordered by the induced subgraph relation is of bounded tree-width if and only if it is subquadratic.*

As we just mentioned, the “only if” part of this theorem readily follows from known facts. The rest of this section will be devoted to the proof of the “if” part, which will be given through a series of intermediate results. In the proof, we use the terminology introduced in Section 6.1. A plan of the proof is outlined in Section 6.2.1. Sections 6.2.2, 6.2.3, 6.2.4, 6.2.5, 6.2.6 contain various parts of the proof.

### 6.2.1 Plan of the proof

To prove Theorem 6.2.1 we will show that graphs with arbitrarily large tree-width contain either arbitrarily large bicliques as subgraphs or arbitrarily large canonical graphs as induced subgraphs. The main notion in our proof is that of a *rake-graph*.

A *rake-graph* (or simply a *rake*) consists of a chordless path, the *base* of the rake, and a number of pendant vertices, called *teeth*, each having a private neighbour on the base. The only neighbour of a tooth on the base will be called the *root* of the tooth, and a rake with  $k$  teeth will be called a  $k$ -rake. We will say that a rake is  $\ell$ -dense if any  $\ell$  consecutive vertices of the base contain at least one root vertex. An example of a 1-dense 9-rake is given in Figure 6.2.

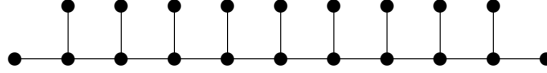


Figure 6.2: 1-dense 9-rake

We will prove Theorem 6.2.1 through a number of intermediate steps as follows.

1. In Section 6.2.2, we observe that any graph of large tree-width contains a rake with many teeth as a subgraph.
2. In Section 6.2.3 we show that any graph containing a rake with many teeth as a subgraph contains either
  - a *dense* rake with many teeth as a subgraph or
  - a large canonical graph as an *induced* subgraph.
3. In Section 6.2.4 we prove that dense rake subgraphs necessarily imply either
  - a large canonical graph as an *induced* subgraph or
  - a large biclique as a subgraph.
4. In Section 6.2.5, we summarize the results of the previous sections to show that any graph of large tree-width contains either
  - a large canonical graph as an *induced* subgraph or
  - a large biclique as a subgraph.
5. In Section 6.2.6, we use the result of Step 4 to prove Theorem 6.2.1.

### 6.2.2 Rake subgraphs in graphs of large tree-width

**Lemma 6.2.1.** *For any natural  $k$ , there is a number  $f(k)$  such that every graph of tree-width at least  $f(k)$  contains a  $k$ -rake as a subgraph.*

*Proof.* A  $k \times k$ -grid is a graph with vertices  $v_{i,j}$   $1 \leq i, j \leq k$  and edges between  $v_{i,j}$  and  $v_{i',j'}$  if and only if  $|i - i'| + |j - j'| = 1$ . In [Robertson et al., 1994], the authors proved that for each  $k$  there is a function  $f(k)$  such that every graph  $G$  of tree-width at least  $f(k)$  has a  $k \times k$ -grid as a minor.

Consequently, any graph  $G$  of tree-width at least  $f(k)$  contains a  $k$ -rake as a minor. It follows that the graph  $G$  contains a subgraph  $H$  from which a  $k$ -rake can be obtained by contraction operations only. We deduce that  $G$  contains a subgraph  $H$ , whose vertices admit a partition  $V(H) = \cup_{i=1}^k V_i \cup_{i=1}^k V'_i$  into disjoint subsets  $V_i$  and  $V'_i$  such that  $G[V_i]$  and  $G[V'_i]$  are connected for each  $i \in \{1, 2, \dots, k\}$ , there is at least one edge with endpoints in both  $V_i$  and  $V_{i+1}$  for each  $i = \{1, 2, \dots, k-1\}$  and there is at least one edge with endpoints in both  $V_i$  and  $V'_i$  for each  $i = \{1, 2, \dots, k\}$ .

To finish the proof we show that the graph  $H$  contains a  $k$ -rake as a subgraph. First, for each  $i = \{1, 2, \dots, k-1\}$ , let  $x_i y_{i+1}$  be an edge with  $x_i \in V_i$  and  $y_{i+1} \in V_{i+1}$ . Then, for each  $i = \{2, 3, \dots, k-1\}$ , as  $G[V_i]$  is connected, we can find a path  $P_i$  in  $G[V_i]$  connecting  $y_i$  and  $x_i$ . We also define  $P_1 = \{x_1\}$  and  $P_k = \{y_k\}$ . These paths will constitute the base of the rake and one can attach tooth  $t_i$  with root in  $P_i$  as follows. If  $V(P_i) = V_i$ , let  $t_i$  be a point in  $V'_i$  which is adjacent to some point in  $V_i$ . Otherwise, if  $V(P_i) \neq V_i$ , let  $t_i$  be a point in  $V_i \setminus V(P_i)$  which has a neighbour in  $V(P_i)$  (possible as  $G[V_i]$  is connected). Thus  $H$ , and hence  $G$ , contains as a subgraph a  $k$ -rake with base  $P_1 \cup P_2 \cup \dots \cup P_k$  and teeth  $\{t_1, t_2, \dots, t_k\}$ . □

### 6.2.3 From rake subgraphs to dense rake subgraphs

The main result of this section is Lemma 6.2.3 below. Its proof is based on the following auxiliary result.

**Lemma 6.2.2.** *Let  $G$  be a graph containing an  $H$ -graph  $H^*$  (possibly tight or semi-tight) as a subgraph with the body being induced (i.e. chordless), and let  $s \geq 2$  an integer. Then*

- (1) *either  $G$  contains a path of length  $t \in \{2, \dots, s+1\}$  connecting a left wing of  $H^*$  to its right wing with all intermediate vertices lying in the body,*
- (2) *or  $G$  contains an induced canonical subgraph of order at least  $s$ .*



*Proof.* Let  $w'$  be a left wing and  $w''$  a right wing of  $H^*$  and  $U = \{u_1, \dots, u_q\}$  its body. Since  $w'$  is adjacent to  $u_1$  and  $w''$  is adjacent to  $u_q$ , there must exist a subpath  $U' = \{u_i, \dots, u_{i+k}\}$  of  $U$  such that  $u_i$  is the only neighbour of  $w'$  in  $U'$  and  $u_{i+t}$  is the only neighbour of  $w''$  in  $U'$ . We assume that  $w', w'', U'$  are chosen so that  $t$  (the length of the path  $U'$ ) is as small as possible. This implies, in particular, that no other wing has an internal neighbor in  $U'$  (i.e. a neighbor different from  $u_i$  and  $u_{i+t}$ ). If  $0 \leq t \leq s-1$ , then  $w'U'w''$  is a path satisfying (1).

Assume now  $k \geq s$ . If  $i = 1$ , we define  $u_{i-1}$  to be the left wing different from  $w'$ , and if  $i+t = q$ , we define  $u_{i+t+1}$  to be the right wing different from  $w''$ . If  $w'$  is adjacent to  $w''$  or  $w'$  is adjacent to  $u_{i+t+1}$  or  $w''$  is adjacent to  $u_{i-1}$ , then a chordless cycle of length at least  $s+1$  arises. Otherwise, the vertices  $w', w'', u_{i-1}, u_i, \dots, u_{i+t}, u_{i+t+1}$  induce a canonical graph of order at least  $s$ .  $\square$

**Lemma 6.2.3.** *Let  $k$  and  $s$  be natural numbers. Every graph containing a  $k+2$ -rake as a subgraph contains either*

- *an  $s+5$ -dense  $k$ -rake a subgraph or*
- *a canonical graph of order at least  $s$  as an induced subgraph.*

*Proof.* Consider a graph  $G$  containing a  $k$ -rake  $R$  as a subgraph. For our construction it is essential that the second and second last vertices of the base of  $R$  are roots while the first and the last vertices are not. To establish this condition we remove the teeth whose roots are the first or the last vertices and possibly shorten the base so that it would start just before the next root and end just after the second last root. This is where  $k+2$  comes from. After this preprocessing, we proceed as follows.

First, we transform any path between any two consecutive root vertices into a shortest, and hence a chordless, path by cutting along any possible chords. Now any two consecutive root vertices together with their teeth, with the path connecting them and with two other their neighbours in the base of  $R$  form an  $H$ -graph satisfying conditions of Lemma 6.2.2. If one of these  $H$ -graphs contains an induced canonical subgraph of order at least  $s$ , the lemma is proved. Therefore, we assume that the wings of each of these graphs are connected by a short path as in (2) of Lemma 6.2.2. We now concatenate (glue) all these paths into the base of a new rake as follows.

Consider three consecutive vertices  $u_{i-1}, u_i, u_{i+1}$  in the base of  $R$  with  $u_i$  being a *not the first* root vertex. Let  $v_i$  be the tooth of  $u_i$ . Also, denote by  $P^l$  a short path connecting two wings of the  $H$ -graph on the left of  $u_i$ , and by  $P^r$  the

respective short path in the  $H$ -graph on the right of  $u_i$ . To simplify the discussion, we will assume that if  $P^r$  starts at  $u_{i-1}$ , then its next vertex is neither  $u_i$  nor  $u_{i+1}$ , since otherwise we can transform  $P^r$  by starting it at  $v_i$ , which will increase the length of the path by at most 1. Also, we will assume that if  $P^r$  starts at  $v_i$ , then its next vertex is not  $u_{i+1}$ , since otherwise we can transform  $P^r$  by adding  $u_i$  between  $v_i$  and  $u_{i+1}$ , which will increase the length of the path by at most 1. We apply similar (symmetric) assumptions with respect to  $P^l$ . With these assumptions in mind, we now do the following.

- If both  $P^l$  and  $P^r$  contain  $u_i$ , then both of them start at  $v_i$  (according to the above assumption). In this case, we glue the two paths at  $u_i$ , define it to be a root vertex in the new rake and define  $v_i$  to be its tooth.
- Assume  $P^l$  contains  $u_i$  (implying it contains  $v_i$ ), while  $P^r$  does not. Assume in addition that  $P_l$  contains  $u_{i-1}$ .
  - If  $P^r$  starts at  $u_{i-1}$ , then we glue the two paths at  $u_{i-1}$  (by cutting  $u_i$  and  $v_i$  off  $P^l$ ), define  $u_{i-1}$  to be a root vertex and  $u_i$  to be its tooth in the new rake.
  - If  $P^r$  starts at  $v_i$ , then we glue the two paths at  $v_i$ , define  $u_i$  to be a root vertex and  $u_{i+1}$  to be its tooth in the new rake.
- The same as in the previous case with the only difference that  $P^l$  does not contain  $u_{i-1}$ ,
  - If  $P^r$  starts at  $u_{i-1}$ , then we replace  $v_i$  by  $u_{i-1}$  in  $P^l$ , glue the two paths at  $u_{i-1}$ , define  $u_i$  to be a root vertex and  $v_i$  to be its tooth in the new rake.
  - If  $P^r$  starts at  $v_i$ , then (like in the previous case) we glue the two paths at  $v_i$ , define  $u_i$  to be a root vertex and  $u_{i+1}$  to be its tooth in the new rake.
- Assume that neither  $P^l$  nor  $P^r$  contains  $u_i$ , then we distinguish between the following cases.
  - If both paths start at  $v_i$ , then we glue them at  $v_i$ , define it to be a root vertex and  $u_i$  its tooth in the new rake.
  - If one of them, say  $P^l$ , starts at  $v_i$ , and the other, i.e.  $P^r$ , at  $u_{i-1}$ , then we concatenate them by adding  $u_i$  (which is adjacent to both  $v_i$  and  $u_{i-1}$ ), define  $u_i$  to be a root vertex and  $u_{i+1}$  its tooth in the new rake.

- If  $P^l$  starts at  $u_{i+1}$   $P^r$  starts at  $u_{i-1}$ , then we again concatenate them by adding  $u_i$ , define  $u_i$  to be a root vertex and  $v_i$  its tooth in the new rake.

The procedure outlined above creates a new rake with  $k$  teeth. The length of each paths used in the construction is initially at most  $s+1$ . In order to incorporate the assumptions regarding  $P^l$  and  $P^r$  we increase them by at most 1 on each end, so the resulting length is at most  $s+3$ . Finally, the process of assignment of roots may require further increase by at most 1 on each end. Hence, we conclude that the new rake is  $s+5$ -dense.  $\square$

#### 6.2.4 Dense rake subgraphs

**Lemma 6.2.4.** *For every  $s, q$  and  $\ell$ , there is a number  $D = D(s, q, \ell)$  such that every graph containing an  $\ell$ -dense  $D$ -rake as a subgraph contains either*

- *a canonical graph of order at least  $s$  as an induced subgraph or*
- *a biclique of order  $q$  as a subgraph.*

*Proof.* To define the number  $D = D(s, q, \ell)$ , we introduce intermediate notations as follows:  $b := 2(q-1)s^q + 2sq + 4$  and  $c := R(2, 2, \max(b, 2q))$ , where  $R$  is the Ramsey number. With these notations the number  $D$  is defined as follows:  $D = D(s, q, \ell) := Z(\ell c^2, 2q, q)$ , where  $Z$  is the number defined in Theorem 6.1.3.

Consider a graph  $G$  containing an  $\ell$ -dense  $D$ -rake  $R^0$  as a subgraph. The base of this rake is a path  $P^0$  of length at least  $D$  and hence, by Theorem 6.1.3, the base contains either a biclique of order at least  $q$  as a subgraph (in which case we are done) or an *induced* path  $P$  of length at least  $\ell c^2$ . Let us call any (inclusionwise) maximal sequence of consecutive vertices of  $P^0$  that belong to  $P$  a *block*. Assume the number of blocks is more than  $c$ . Take the first  $c$  of them and add to each of them the vertex of  $P^0$  following the rightmost endpoint of the block as a tooth. This creates a  $c$ -rake with  $P$  being an induced base. If the number of blocks is at most  $c$ , then  $P^0$  must contain a block of size at least  $\ell c$ , in which case this block also forms an induced base of a  $c$ -rake (since  $R^0$  is  $\ell$ -dense). We see that in either case  $G$  has a  $c$ -rake with an induced base. According to the definition of  $c$ , the  $c$  teeth of this rake induce a graph which has either a clique of size  $2q$  (and hence a biclique of order  $q$  in which case we are done), or an independent set of size  $b$ . By ignoring the teeth outside this set we obtain a  $b$ -rake  $R$  with an induced base and with teeth forming an independent set.

Let us denote the base of  $R$  by  $U$ , its vertices by  $u_1, \dots, u_m$  (in the order of their appearances in the path), and the teeth of  $R$  by  $t_1, \dots, t_b$  (following the order

of their root vertices).

Denote  $r := (q - 1)s^q + 2$  and consider two sets of teeth  $T_1 = \{t_2, t_3, \dots, t_r\}$  and  $T_2 = \{t_{b-1}, t_{b-2}, \dots, t_{b-r+1}\}$ . By definition of  $r$  and  $b$ , there are  $2sq$  other teeth between  $t_r$  and  $t_{b-r+1}$ , and hence there is a set  $M$  of  $2sq$  consecutive vertices of  $U$  between the root of  $t_r$  and the root of  $t_{b-r+1}$ . We partition  $M$  into  $2q$  subsets (of consecutive vertices of  $U$ ) of size  $s$  each and for  $i = 1, \dots, 2q$  denote the  $i$ -th subset by  $M_i$ .

If each vertex of  $T_1$  has a neighbour in each of the first  $q$  sets  $M_i$ , then by the Pigeonhole Principle there is a biclique of order  $q$  with  $q$  vertices in  $T_1$  and  $q$  vertices in  $M$  (which can be proved by analogy with Lemma 6.1.1). Similarly, a biclique of order  $q$  arises if each vertex of  $T_2$  has a neighbour in each of the last  $q$  sets  $M_i$ . Therefore, we assume that there are two vertices  $t_a \in T_1$  and  $t_b \in T_2$  and two sets  $M_x$  and  $M_y$  with  $x < y$  such that  $t_a$  has no neighbors in  $M_x$ , while  $t_b$  has no neighbours in  $M_y$ .

By definition,  $t_a$  has a neighbour in  $U$  (its root) on the left of  $M_x$ . If additionally  $t_a$  has a neighbour to the right of  $M_x$ , then a chordless cycle of length at least  $s$  arises (since  $|M_x| = s$  and  $t_a$  has no neighbors in  $M_x$ ), in which case the lemma is true. This restricts us to the case, when all neighbours of  $t_a$  in  $U$  are located to the left of  $M_x$ . By analogy, we assume that all neighbours of  $t_b$  in  $U$  are located to the right of  $M_y$ . Let  $u_i$  be the rightmost neighbour of  $t_a$  in  $U$  and  $u_j$  be the leftmost neighbour of  $t_b$  in  $U$ . According to the above discussion,  $i < j$  and  $j - i > 2s$ . But then the vertices  $t_a, t_b, u_{i-1}, u_i, \dots, u_j, u_{j+1}$  induce an  $H$ -graph (possibly tight or semi-tight) of order more than  $s$  (the existence of vertices  $u_{i-1}$  and  $u_{j+1}$  follows from the fact that  $T_1$  does not include  $t_1$ , while  $T_2$  does not include  $t_b$ ).  $\square$

### 6.2.5 Canonical graphs and bicliques in graphs of large tree-width

**Theorem 6.2.2.** *For every  $s, q$ , there is a number  $X = X(s, q)$  such that every graph of tree-width at least  $X$  contains either*

- *a canonical graph of order at least  $s$  as an induced subgraph or*
- *a biclique of order  $q$  as a subgraph.*

*Proof.* We define  $X(s, q)$  as  $X(s, q) := f(D(s, q, s + 5) + 2)$ , where  $f$  comes from Lemma 6.2.1 and  $D$  comes from Lemma 6.2.4. If a graph  $G$  has tree-width at least  $X(s, q)$ , then by Lemma 6.2.1 it contains  $D(s, q, s + 5) + 2$ -rake  $R$  as a subgraph. Then, by Lemma 6.2.3,  $G$  contains either a canonical graph of order at least  $s$  as an

induced subgraph, or an  $s + 5$ -dense  $D(s, q, s + 5)$ -rake as a subgraph. In the first case, the theorem is proved. In the second case, we conclude by Lemma 6.2.4 that  $G$  contains either a canonical graph of order at least  $s$  as an induced subgraph or a biclique of order  $q$  as a subgraph.  $\square$

### 6.2.6 Proof of Theorem 6.2.1

Let  $\mathcal{Y}$  be a hereditary class of graphs defined by a set of forbidden induced subgraphs  $M$ , and assume  $\mathcal{Y}$  is well-quasi-ordered by the induced subgraph relation.

If  $\mathcal{Y}$  is of bounded tree-width, then the set  $M$  of forbidden graphs must include a complete graph and a complete bipartite graph, since the tree-width is bounded neither for complete graphs nor for complete bipartite graphs. Therefore,  $\mathcal{Y}$  is subquadratic, which proves the first part of the theorem.

Assume now that  $\mathcal{Y}$  is subquadratic. Then there must exist  $q$  such that no graph in  $\mathcal{Y}$  contains  $K_{q,q}$  as a subgraph. Suppose by contradiction that  $\mathcal{Y}$  contains an infinite sequence  $\mathcal{Y}'$  of graphs of increasing tree-width. In this sequence, there must exist a graph  $G^1$  of tree-width at least  $X(s, q)$ , where  $X(s, q)$  is defined in Theorem 6.2.2 and  $s$  is an arbitrarily chosen constant. Then by Theorem 6.2.2  $G^1$  contains a canonical graph  $H^1$  of order at least  $s$ . We denote the order of  $H^1$  by  $s_1$  and find in  $\mathcal{Y}'$  a graph  $G^2$  of tree-width at least  $X(s_1 + 1, q)$ .  $G^2$  must contain a canonical graph  $H^2$  of order  $s_2 \geq s_1 + 1$ , and so on. In this way, we construct an infinite sequence  $H^1, H^2, \dots$ , which form an antichain by Claim 6.1.1. This contradicts the assumption that  $\mathcal{Y}$  is well-quasi-ordered by the induced subgraph relation and hence shows that  $\mathcal{Y}$  is of bounded tree-width.

## 6.3 Deciding WQO for subquadratic classes

Recall our fundamental problem of deciding well-quasi-ordering for graphs: given a finite collection of graphs  $G_1, \dots, G_k$ , decide whether the class  $\text{Free}(G_1, \dots, G_k)$  is well-quasi-ordered by induced subgraphs or not. In this section we remark that for subquadratic properties the problem has a simple solution.

**Theorem 6.3.1.** *Let  $X$  be a subquadratic class of graphs defined by a finite collection  $F$  of forbidden induced subgraphs. Then  $X$  is well-quasi-ordered by the induced subgraph relation if and only if  $F$  contains a path  $P_k$ .*

*Proof.* If  $F$  does not contain a path  $P_k$ , then  $X$  contains infinitely many cycles, because by forbidding finitely many graphs we can exclude only finitely many cycles from  $X$ .

Assume now that  $F$  contains a path  $P_k$ . Since  $X$  is subquadratic,  $F$  must also contain a clique and a biclique. But then, by Theorem 6.1.3, graphs in  $X$  do not contain large paths as subgraphs. In other words, there is a constant  $t$  such that  $X$  is a subclass of the class  $Y$  containing no  $P_t$  as a subgraph. In [Ding, 1992], the author showed that a class of graphs closed under taking subgraphs is well-quasi-ordered by the induced subgraph relation if and only if it contains finitely many cycles and finitely many graphs of the form  $H_i$ . According to this result,  $Y$ , and hence  $X$ , is well-quasi-ordered by induced subgraphs.  $\square$

## 6.4 Byproduct: Computing small bicliques in graphs without long induced paths

Given a graph  $G$  and parameter  $k$ , the Biclique problem asks if  $G$  has a biclique of order  $k$  as a subgraph (not necessarily induced). This problem appeared under the name of 'Balanced Complete Bipartite Subgraph' as problem [GT24] in the famous book of Garey and Johnson, an NP-hardness proof has been further provided by Johnson [1987]. An application of the problem to the VLSI design is described in detail in [Arbib and Mosca, 1999], where the authors also develop polynomial time algorithms for a number of restricted classes of graphs. Exact exponential time algorithms for this problem can be found in [Binkele-Raible et al., 2010]. The related *induced* Biclique problem is known to be W[1]-hard [Chen et al., 2008]. However, the parameterized complexity of the Biclique problem is a longstanding open question that received significant attention from the parameterized complexity community and is believed to be W[1]-hard (see the abstract of [Bulatov and Marx, 2014]).

In this section, we consider a restricted version of this problem by introducing an additional parameter  $s$  and assuming  $G$  to be  $P_s$ -free, i.e. without induced paths of length  $s$ . This restriction can be formally stated as follows (the abbreviation NLIP in the name of this problem stands for 'No Long Induced Paths').

NLIP-BICLIQUE

*Input:* A graph  $G$

*Parameters:*  $k, s$

*Assumption:*  $G$  is  $P_s$ -free

*Output:* A biclique of  $G$  of size at least  $k$  or 'NO' if there is no such biclique.

We show that under this additional parameterization the biclique problem becomes fixed-parameter linear. Let us remark that the parameterization by  $s$  alone

is not enough for efficient computation of a largest biclique, because this problem is NP-hard on  $P_s$ -free graphs for  $s \geq 8$ , as we observe below.

Graphs without long induced paths, i.e.  $P_s$ -free graphs for a constant  $s$ , have been extensively studied in the literature (see e.g. [Bacsó and Tuza, 1990, Dong, 1996, Lozin and Rautenbach, 2003]). For small values of  $s$ , the structure of  $P_s$ -free graphs is simple. For instance,  $P_3$ -free graphs are precisely the graphs every connected component of which is a clique.  $P_4$ -free graphs also enjoy many nice properties. In particular, the clique-width of  $P_4$ -free graphs is bounded by a constant and hence many algorithmic problems that are generally NP-hard admit polynomial time solutions when restricted to  $P_4$ -free graphs.

In the class of  $P_s$ -free graphs with  $s \geq 5$ , the situation changes drastically and the computational complexity changes from polynomial-time solvability to NP-hardness for many important algorithmic graph problems. For instance, VERTEX COLORING [Král et al.] and MINIMUM DOMINATING SET [Korobitsyn, 1990] are NP-hard for  $P_5$ -free graphs, and VERTEX 4-COLOURABILITY is NP-hard for  $P_8$ -free graphs [Broersma et al., 2012]. For many other problems, the complexity status on graphs without long induced paths is unknown. For instance, the complexity status is unknown for MAXIMUM INDEPENDENT SET in  $P_s$ -free graphs with  $s \geq 5$  and for VERTEX 3-COLOURABILITY in  $P_s$ -free graphs with  $s \geq 7$  (for some partial results related to these problems we refer the reader to [Le et al., 2007, Randerath and Schiermeyer, 2004, Woeginger et al., 2001]).

Now we extend this series of negative results by showing that the Biclique problem is NP-hard on  $P_s$ -free graphs for  $s \geq 8$ . This justifies the use of  $s$  as an additional parameter.

**Proposition 6.4.1.** *Computing maximum biclique in a  $P_s$ -free graph is NP-hard for  $s \geq 8$*

*Proof.* This is implicitly proven in [Johnson, 1987] because an instance of the Clique problem is reduced to an instance of the Biclique problem on a  $P_8$ -free graph. Indeed, by construction, given a graph  $G$ , a bipartite graph  $H$  is constructed in which the first part  $A$  corresponds to the edges of  $G$  and the second part  $B$  corresponds to a superset of its vertices and each vertex of  $A$  is adjacent to all vertices of  $B$  but those corresponding to the endpoints of the respective edge. It is not hard to see that  $H$  is  $P_8$ -free. Indeed, let  $P$  be a path of length 8. It has one terminal vertex  $u$  in  $A$ , and 4 vertices  $v_1, \dots, v_4$  of  $B$  included in it. Three vertices out of  $v_1, \dots, v_4$  are non-adjacent to  $u$  in  $P$  but only two of them may be the endpoints of the respective edge. It follows that  $u$  is necessarily adjacent in  $H$  to the remaining one, thus producing a chord in  $P$ .  $\square$

The proof of fixed-parameter tractability of the NLIP-BICLIQUE problem is based on Theorem 6.1.3, which implies the following corollary.

**Corollary 6.4.1.** *For any natural numbers  $s$  and  $k$  there is a natural number  $T(s, k)$  such that any graph of treewidth at least  $T(s, k)$  either has an induced path of length  $s$  or a biclique of order  $k$ .*

*Proof.* It is well known (see e.g. Theorem 9 of [Fellows and Langston, 1989]) that for each natural  $r$  there is  $Y(r)$  such that if the treewidth of the given graph is at least  $Y(r)$ , the graph has a path of size at least  $r$ . Take  $T(s, k) = Y(Z(s, 2k, k))$  and apply Theorem 6.1.3.  $\square$

**Theorem 6.4.1.** *For fixed parameters  $s$  and  $k$ , the NLIP-BICLIQUE problem can be solved in a linear time.*

*Proof.* Let  $G$  be the input graph with  $n$  vertices. Using the linear time algorithm of Bodlaender [1993], test the existence of a path of length  $Z(s, 2k, k)$  and find it, in case it exists.

Assume that such a path  $P$  has been found. In this case, the subgraph of  $G$  induced by the vertices of  $P$  has a biclique of size  $k$  as follows from Theorem 6.1.3. Since the size of this subgraph depends only on the parameters, the way this biclique is computed does not affect the desired runtime so, we can use the brute force.

If  $G$  does not have a path of length  $Z(s, 2k, k)$  then according to Corollary 6.4.1, the treewidth is at most  $T(s, q)$ , therefore, the biclique problem can be solved by standard techniques for graphs of bounded treewidth, say Courcelle's theorem [Courcelle et al., 2000].  $\square$

## 6.5 Conclusion

In this section we verified the conjecture of Daligault et al. [2010] for a family of subquadratic properties and showed that the question of deciding well-quasi-ordering can be done in linear time for finitely defined subquadratic properties. Answering these questions for hereditary classes containing arbitrarily large cliques or bicliques remains a challenging research question.

Our approach was based on Ramsey-type result, Theorem 6.1.3, which states that for every  $s, t$  and  $q$  there is a number  $Z = Z(s, t, q)$  such that every graph with a path of length at least  $Z$  contains either  $P_s$  or  $K_t$  or  $K_{q,q}$  as an induced subgraph. Our result proves the existence of the number  $Z(s, t, q)$  but does not tell anything



about its value. We derived some bounds for the case when  $t = 3, q = 2$ . Deriving any reasonable bounds in general setting is an interesting and challenging research question.

It would be also interesting to see how Theorem 6.1.3 could be modified and extended. Our result is equivalent to saying that in a subquadratic property any 'large' graph contains a long induced path. Here we measure largeness by size of a long (not necessarily induced) path in the graph. If instead we replaced this condition by saying that a graph is 'large' if it has a large average degree, then by the result of Kühn and Osthus [2004] we could claim that our 'large' graph contains an induced subdivision of any graph. One may ask, what happens if the largeness condition is replaced by large tree-width, large clique-width, ect. Can we claim existence of a long induced path? Can we claim anything stronger than that?

## Chapter 7

# Classes of unbounded clique-width

Daligault et al. [2010] conjecture states that every well-quasi-ordered class has bounded clique-width. It is not hard to see that this conjecture is equivalent to the statement that every minimal class of unbounded clique-width is not well-quasi-ordered. We conjecture that even a stronger relation holds: each minimal class of unbounded clique-width contains a canonical antichain. Lozin [2011] identified the first two minimal hereditary classes of unbounded clique-width: bipartite permutation graphs and unit interval graphs. Further, the existence of canonical antichains were shown for these classes, in [Lozin and Mayhill, 2011].

Recently, [Korpelainen et al., 2014] identified another class of unbounded clique-width: split permutation graphs. In this chapter we prove that this class and its bipartite analog – class of bichain graphs are indeed minimal classes of unbounded clique-width. In Section 7.1 we obtain a preliminary result: we construct universal graphs for these classes. As a byproduct, we also construct a proper universal 321-avoiding permutation. In Section 7.2 we use the universal construction for graphs to prove that both the class of split permutation graphs and the class of bichain graphs are indeed minimal classes of unbounded clique-width. In section 7.3 we prove that each of these classes contains a canonical antichain with respect to labelled induced subgraph relation. The restriction to the labelled case could be partly justified by Conjecture 3.1.1 and the fact that these two classes are finitely defined (proved in [Korpelainen, 2012]).

## 7.1 Universal graphs and universal permutations

Let  $X$  be a family of graphs and  $X_n$  the set of  $n$ -vertex graphs in  $X$ . A graph containing all graphs from  $X_n$  as induced subgraphs is called  $n$ -universal for  $X$ . The problem of constructing universal graphs is closely related to graph representations and finds applications in theoretical computer science [Alstrup and Rauhe, 2002, Kannan et al., 1992]. This problem is trivial if universality is the only requirement, since the union of all vertex disjoint graphs from  $X_n$  is obviously  $n$ -universal for  $X$ . However, this construction is generally neither optimal, in terms of the number of its vertices, nor proper, in the sense that it does not necessarily belong to  $X$ .

Let us denote an  $n$ -universal graph for  $X$  by  $U^{(n)}$  and the set of its vertices by  $V(U^{(n)})$ . Since the number of  $n$ -vertex subsets of  $V(U^{(n)})$  cannot be smaller than the number of graphs in  $X_n$ , we conclude that

$$\log_2 |X_n| \leq \log_2 \binom{|V(U^{(n)})|}{n} \leq n \log_2 |V(U^{(n)})|.$$

Also, trivially,  $n \leq |V(U^{(n)})|$ , and hence,

$$n \log_2 n \leq n \log_2 |V(U^{(n)})|.$$

We say that  $U^{(n)}$  is *optimal* if  $n \log_2 |V(U^{(n)})| = \max(\log_2 |X_n|, n \log_2 n)$ , *asymptotically optimal* if

$$\lim_{n \rightarrow \infty} \frac{n \log_2 |V(U^{(n)})|}{\max(\log_2 |X_n|, n \log_2 n)} = 1,$$

and *order-optimal* if there is a constant  $c$  such that for all  $n \geq 1$ ,

$$\frac{n \log_2 |V(U^{(n)})|}{\max(\log_2 |X_n|, n \log_2 n)} \leq c.$$

Optimal universal graphs (of various degrees of optimality) have been constructed for many graph classes such as the class of all graphs [Moon, 1965], threshold graphs [Hammer and Kelmans, 1994], planar graphs [Chung, 1990], graphs of bounded arboricity [Alstrup and Rauhe, 2002], of bounded vertex degree [Butler, 2009, Esperet et al., 2008], split graphs and bipartite graphs [Lozin, 1997], bipartite permutation graphs [Lozin and Rudolf, 2007], etc. Some of these constructions can also be extended to an infinite universal graph, in which case the question of optimality is not relevant any more and the main problem is finding a universal element *within* the class under consideration. We call a universal graph for a class  $X$  that belongs to  $X$  a *proper* universal graph.

Not for every class there exist proper infinite universal graphs with countably many vertices [Füredi and Komjáth, 1997]. Such constructions are known for the class of all graphs [Rado, 1964],  $K_4$ -free graphs and some other classes [Komjáth et al., 1988, Komjáth and Pach, 1984]. A proper infinite countable universal graph can be also easily constructed for the class of bipartite permutation graphs from finite  $n$ -universal graphs (represented in Figure 7.1) by increasing  $n$  to infinity.

In the present section, we study a related class, namely, the class of split permutation graphs. In spite of its close relationship to bipartite permutation graphs, finding a *proper*  $n$ -universal graph for this class is not a simple problem even for finite values of  $n$ . We solve this problem by establishing a bijection between symmetric split permutation graphs and 321-avoiding permutations and by constructing a proper universal 321-avoiding permutation. Our construction uses  $2n^3$  vertices. Since there are at most  $n! < n^n$  labelled permutation graphs, this construction is order-optimal. Whether this construction can be extended to an infinite countable graph remains a challenging open problem.

In Section 7.1.1 we construct a universal 321-avoiding permutation and in Section 7.1.2 we use this construction to build a universal split permutation graph. For all definitions concerning permutations and permutation graphs we refer reader to Section 4.1.

### 7.1.1 Universal 321-avoiding permutations

In this section, we study the set of 321-avoiding permutations, i.e. permutations containing no 321 as a pattern. In other words, a permutation is 321-avoiding if it contains no subsequence of length 3 in decreasing order reading from left to right, or equivalently, if its elements can be partitioned into at most 2 increasing subsequences. For example, the permutation 245136 avoids the pattern 321 (with 2456 and 13 being two increasing subsequences), while 261435 does not because of the subsequence 643. It is known that the number of 321-avoiding permutations of length  $n$  is given by  $C_n = \frac{1}{n+1} \binom{2n}{n}$ , the  $n$ -th *Catalan number*. We let  $S_n(321)$  denote the set of all 321-avoiding permutations of length  $n$ .

The set of 321-avoiding permutations have been studied regularly in the *theory of permutation patterns* (see [Kitaev, 2011] for a recent comprehensive introduction to the respective field) in connection with various combinatorial problems. In this section, we study these permutations in connection with the notion of a *universal permutation*.

Given a set  $X$  of permutations, we say that a permutation  $\Pi$  is  $n$ -universal for  $X$  if it contains all permutations of length  $n$  from  $X$  as patterns. Moreover,  $\Pi$

is a *proper*  $n$ -universal permutation for  $X$  if  $\Pi$  belongs to  $X$ .

Note that it is straightforward to construct a proper  $n$ -universal permutation for  $S_n(321)$  of length  $nC_n = \frac{n}{n+1} \binom{2n}{n}$ . Indeed, we can list all the  $C_n$  permutations in a row, say, in a lexicographic order, and create a single permutation by raising the elements of the  $i$ -th permutation from left by  $(i-1)n$ , for  $1 \leq i \leq C_n$ ; the resulting permutation will clearly avoid the pattern 321. For example, for  $n = 3$ , the elements of  $S_3(321)$  can be listed as 123, 132, 213, 231 and 312 leading to the permutation

$$123465879(11)(12)(10)(15)(13)(14).$$

However, our goal in this section is to construct a proper  $n$ -universal permutation of length  $n^2$  for the set of 321-avoiding permutations of length  $n$ . Our construction is suggested by the following two facts:

- the permutation graphs of 321-avoiding permutations are bipartite permutation graphs,
- the graph  $X_{n,n}$  represented in Figure 7.1 is  $n$ -universal for the class of bipartite permutation graphs [Lozin and Rudolf, 2007].

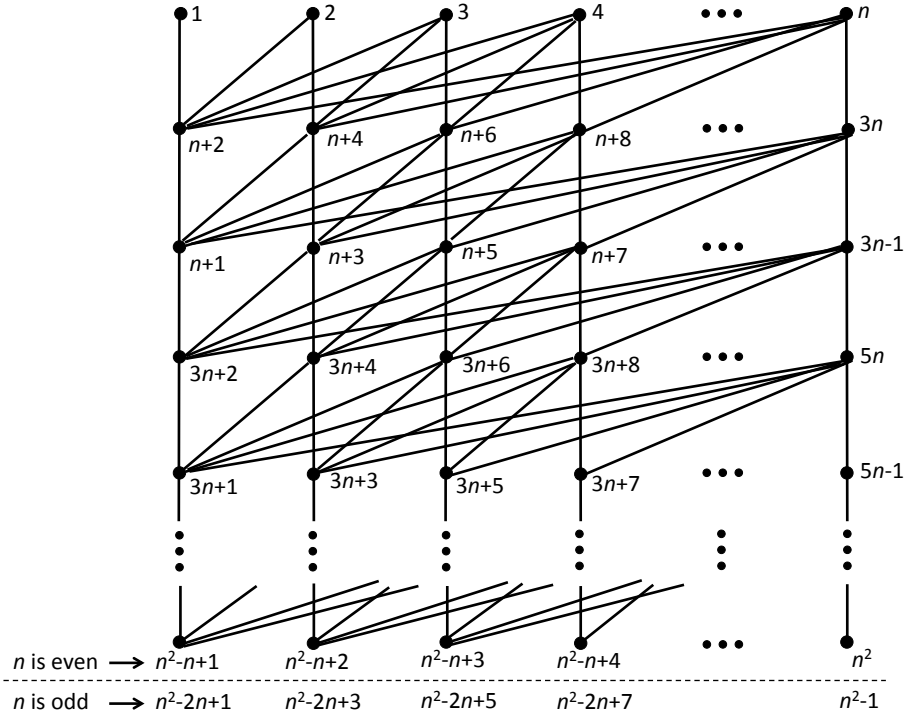


Figure 7.1: The universal bipartite permutation graph  $X_{n,n}$  with a labeling.

In order to construct an  $n$ -universal permutation for the class of 321-avoiding permutations, we label the graph  $X_{n,n}$  in a specific way and define a permutation corresponding to that labeling. The labeling is shown in Figure 7.1 and the corresponding permutation is shown in Figure 7.2. We denote this permutation  $\rho_n$ . It begins with  $(n+2)1(n+4)2 \dots (3n)n$  followed by  $(3n+2)(n+1)(3n+4)(n+3) \dots (5n)(3n-1)$  followed by  $(5n+2)(3n+1)(5n+4)(3n+3) \dots (7n)(5n-1)$  followed by  $(7n+2)(5n+1)(7n+4)(5n+3) \dots (9n)(7n-1)$ , etc. In general, for  $1 < i < \lfloor \frac{n}{2} \rfloor$ , the  $i$ -th  $2n$ -block of the permutation is given by

$$(2ni - n + 2)(2ni - 3n + 1)(2ni - n + 4)(2ni - 3n + 3) \dots (2ni + n)(2ni - n - 1).$$

For  $n \geq 3$ , in case of even  $n$ , the last  $2n$  elements of the permutation are

$$(n^2 - n + 1)(n^2 - 3n + 1)(n^2 - n + 2)(n^2 - 3n + 3) \dots (n^2)(n^2 - n - 1)$$

while in case of odd  $n$ , the last  $n$  elements of the permutation are

$$(n^2 - 2n + 1)(n^2 - 2n + 3) \dots (n^2 - 1).$$

For small values of  $n$ , we have the following:

$n$	$\rho_n$
1	1
2	3142
3	517293468
4	6182(10)3(12)4(13)5(14)7(15)9(16)(11)
5	7192(11)3(13)4(15)5(17)6(19)8(21)(10)(23)(12)(25)(14)(16)(18)(20)(22)(24)

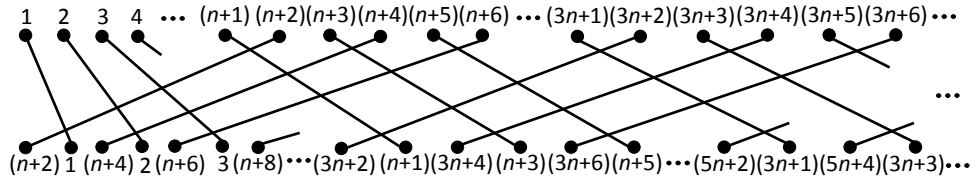


Figure 7.2: The permutation  $\rho_n$  corresponding to the graph in Figure 7.1.

If in a 321-avoiding permutation of length  $n$  the element  $n-1$  is to the right of  $n$ , then  $n-1$  must be the rightmost element in the permutation to avoid the pattern 321 involving the elements  $n-1$  and  $n$ . Using this observation, we can generate all 321-avoiding permutations of length  $n$  from 321-avoiding permutations

of length  $n-1$  as follows. Let  $\pi \in S_{n-1}(321)$ . To obtain all permutations in  $S_n(321)$  derived from  $\pi$ , we can insert  $n$  in  $\pi$  in any place to the right of  $n-1$  (this cannot lead to an occurrence of the pattern 321, say  $nxy$ , since  $(n-1)xy$  would then be an occurrence of 321 in  $\pi$ ); also, we can replace the element  $(n-1)$  in  $\pi$  by  $n$  and adjoin the element  $(n-1)$  to the right of the obtained permutation. Clearly, if we apply the described operations to different permutations in  $S_{n-1}(321)$ , we will obtain different permutations in  $S_n(321)$ . Moreover, the steps described above are reversible, namely for any permutation in  $S_n(321)$  we can figure out from which permutation in  $S_{n-1}(321)$  it was obtained, which gives us the desired.

The main result of the section is the following theorem.

**Theorem 7.1.1.** *The permutation  $\rho_n$  is a proper  $n$ -universal permutation for the set  $S_n(321)$ .*

*Proof.* Note that if one removes the last column and last row in the graph in Figure 7.1 and relabels vertices by letting a vertex receive label  $i$  if it is the  $i$ -th largest label in the original labeling, then one gets exactly the graph in question of size  $(n-1) \times (n-1)$ . This observation allows us to apply induction on the size of the graph with obvious base case on 1 vertex giving a permutation containing the only 321-avoiding permutation of size 1.

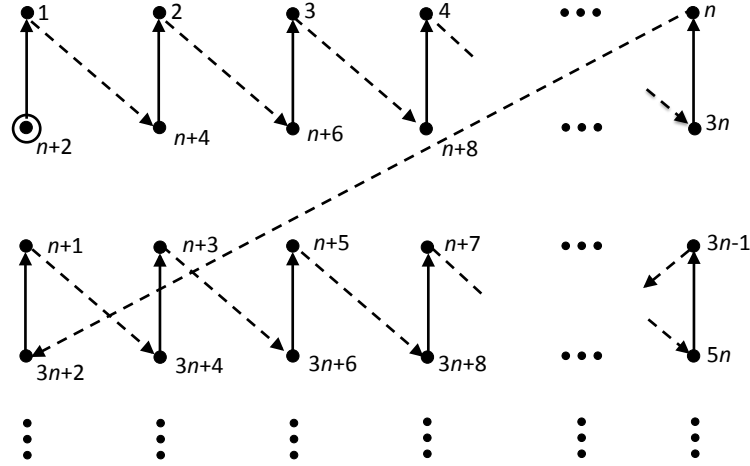


Figure 7.3: A direct reading off the permutation  $\rho_n$  from the graph.

It is not hard to see based on Figure 7.2 that a direct way to read off the permutation corresponding to the graph in Figure 7.1 is to start with the element  $n+2$  and to follow the (solid and dashed) arrows as shown in Figure 7.3; in the case of odd  $n$ , the last row only has horizontal arrows (say, solid) going from left to right as shown schematically in the left picture in Figure 7.4. We only pay attention

to the vertical solid arrows and introduce some terminology here. We say that an element is in a *top row* if an arrow points at it. By definition, all elements in the last row for the case when  $n$  is odd are considered to be on a top row. Non-top elements are said to be from a *bottom row*. Thus, each element is either a top or a bottom element depending on which row it lies on. Two elements are said to be *neighbor elements* if they are connected by a vertical arrow. Clearly, for two neighbors we always have that the bottom neighbor is larger than the top neighbor. Also, it is straightforward to see from the structure of the graph that on the same row, elements increase from left to right. Finally, a bottom element is always larger than the top element right below it.

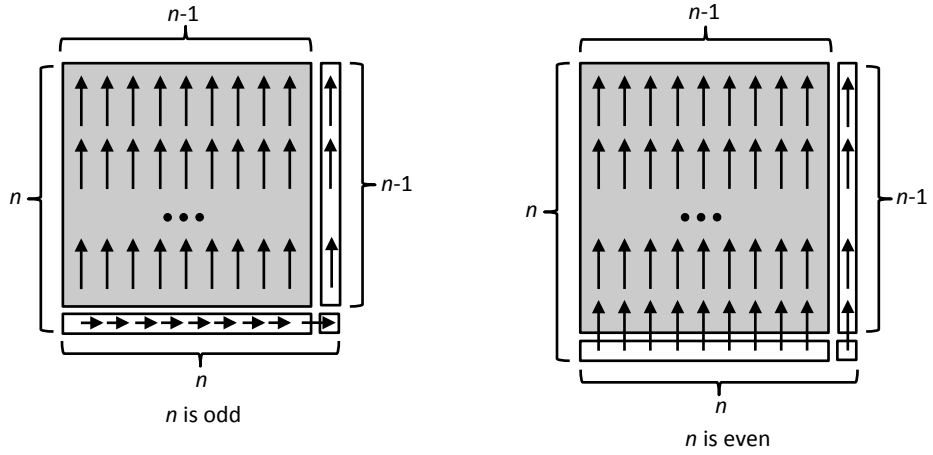


Figure 7.4: Applying inductive hypothesis for Theorem 7.1.1

Suppose that the statement is true for the case of  $n - 1$  and we would like to prove it for the case of  $n$ . That is, if in the graph on  $n^2$  vertices we remove the last row and the last column, then in the obtained graph, shown schematically shaded in Figure 7.4, we can *realize* any 321-avoiding permutation of length  $n - 1$ . By realization of a permutation of length  $n - 1$  we mean picking  $n - 1$  elements in the shaded area in Figure 7.4, so that while going through the directed path in Figure 7.3, the picked elements form the same relative order as the elements of the permutation. Our goal is to show how to realize a 321-avoiding permutation  $\pi$  of length  $n$  based on a realization of the permutation  $\sigma$  obtained from  $\pi$  by removing the largest element. Note that the element  $n^2$  is not in the shaded area. We distinguish here between two cases.

*Case 1:  $n$  is odd.* We split this case into two subcases as follows.

*Case 1.1:  $\pi$  does not end on  $n - 1$ .* That is, in  $\pi$ ,  $n - 1$  is to the left of  $n$ . If



$\sigma$  ends with the largest element, we can always take the element  $n^2$  to extend the realization of  $\sigma$  to a realization of  $\pi$ . On the other hand, if  $n - 1$  is not the rightmost element in  $\sigma$ , then clearly the element corresponding to  $n - 1$ , say  $x$ , must be in a bottom row, say, row  $i$  ( $i$  is even). Moreover, since the elements to the right of  $n - 1$  in  $\sigma$  must be in increasing order to avoid the pattern 321, and we have at most  $n - 2$  such elements, we know that those elements are chosen among the top elements in row  $i - 1$  weakly to the right of  $x$  and, if  $i + 1 < n - 1$ , possibly among the top elements in row  $i + 1$  weakly to the left of  $x$ . In either case, we can always extend the realization of  $\sigma$  to a realization of  $\pi$  by picking a bottom element to the right of  $x$  in row  $i$  or weakly to the left of  $x$  in row  $i + 2$  (if it exists). Indeed, if  $n$  is next to the left of an element  $y$  in  $\pi$ , and vertex  $z$  corresponds to  $y$  in the realization of  $\sigma$ , then we can pick the (bottom) neighbor of  $z$  to correspond to  $n$  and thus to get a realization of  $\pi$ . For example, if  $\pi = (24513)$  ( $n = 5$ ) and the realization of  $\sigma = (2413)$  is the sequence 7, 11, 6, 10, that is,  $x = 11$ ,  $y = 1$  and  $z = 6$ , then the neighbor of  $z$  is 17 and a desired realization of  $\pi$  is 7, 11, 17, 6, 10; see the left picture in Figure 7.5.

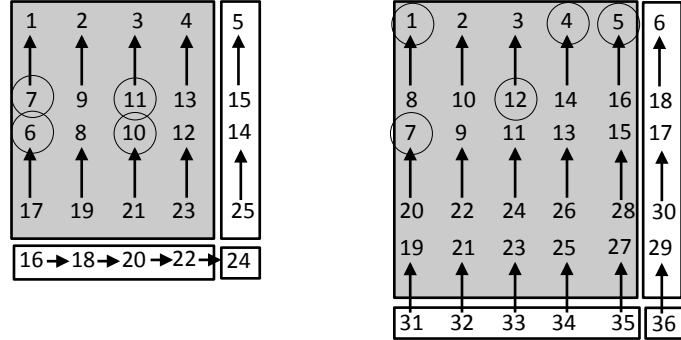


Figure 7.5: Examples illustrating our inductive proof of Theorem 7.1.1

The only case when the described approach does not work is if the bottom neighbor of  $z$  is  $x$  itself. In this case we will take a different realization of  $\sigma$  obtained by one position shift to the right of all the elements already involved in the realization that are located weakly above or weakly to the right of  $x$  including  $x$  itself, and we pick the element directly below the shifted  $z$  to obtain the desired outcome. See the respective subcase involving shifting in Case 1.2 below for more details on why the shift works and for an example of applying the shift.

*Case 1.2:  $\pi$  ends on  $n - 1$ .* If the element  $x$  corresponding to  $n - 1$  in a realization of  $\sigma$  is in a top row,  $\sigma$  must end with  $n - 1$  and we can add the (bottom) neighbor of  $x$  to the realization of  $\sigma$  to get a realization of  $\pi$  (then the neighbor of  $x$  will

correspond to  $n$ , while the other elements in  $\pi$  will have the same corresponding elements as those in  $\sigma$ ). On the other hand, if  $x$  is in a bottom row, say  $i$ , then applying the same arguments as above, everything to the right of  $n - 1$  in  $\sigma$  can be realized by the elements weakly to the right of  $x$  in (top) row  $i - 1$  and weakly to the left of  $x$  in (top) row  $i + 1$  if  $i + 1 < n - 1$ .

The worst case for us here is if the element directly below  $x$  is involved in the realization of  $\sigma$ . In this case we will take a different realization of  $\sigma$  obtained by one position shift to the right of all the elements already involved in the realization that are located weakly above or weakly to the right of  $x$  including  $x$  itself (that is, we shift the realization elements in the first quadrant, including the positive semi-axes, with the origin at  $x$ ). Such a shift is always possible because we have the  $n$ -th column adjoined. Also, such a shift does not change any order of elements involved in the initial realization of  $\sigma$  because there are no elements of the realization in the fourth quadrant and the positive  $x$ -axis with the origin at  $x$ . For example, if  $n = 5$ ,  $x = 11$  and 10 and, say 3 and 4 would be in a realization of  $\sigma$ , after the shift, we would have 10,  $x = 13$ , 4 and 5 in the realization of  $\sigma$  instead of 10, 11, 3 and 4, respectively (see the left picture in Figure 7.5 for layout of elements when  $n = 5$ ; the circles there should be ignored for the moment).

In either case, we can pick the element directly below  $x$  to play the role of  $n - 1$ , while  $x$  will be playing the role of  $n$  in the obtained realization of  $\pi$ . For example, if  $\pi = (25134)$  and a realization of  $\sigma = (2413)$  is the sequence 7, 11, 6, 10 as in our previous example ( $x = 11$  here), then we can consider instead the realization of  $\sigma$  7, 13, 6, 10 and 7, 13, 6, 10, 12 is a desired realization of  $\pi$ ; see the left picture in Figure 7.5.

*Case 2:  $n$  is even.*

In this case, we actually can repeat verbatim all the arguments from the case “ $n$  is odd” keeping in mind that

- by definition, the last row is a top row, and
- picking the bottom neighbor element may involve the last row, which does not change anything.

Of course, different examples should be provided in this case. For the subcase “ $\pi$  does not end on  $n - 1$ ”, if  $\pi = 152634$  ( $n = 6$ ) and the realization of  $\sigma = 15234$  is the sequence 1, 12, 4, 5, 7, that is,  $x = 12$ ,  $y = 3$  and  $z = 5$ , then the neighbor of  $z$  is 16 and the desired realization of  $\pi$  is 1, 12, 4, 16, 5, 7; see the right picture in Figure 7.5. Finally, for the subcase “ $\pi$  ends on  $n - 1$ ”, if  $\pi = 162345$  and a

realization of  $\sigma = 15234$  is the sequence 1, 12, 4, 5, 7 ( $x = 12$ ), then 1, 12, 4, 5, 7, 11 is a desired realization of  $\pi$ ; see the right picture in Figure 7.5.  $\square$

### 7.1.2 Universal split permutation and universal bichain graphs

In this section, we use the result of Section 7.1.1 to construct a universal split permutation and universal bichain graphs. To this end, let us first introduce some more terminology.

A *vicinal quasi-order*  $\sqsubseteq$  on the vertex set of a graph is defined as:

$$x \sqsubseteq y \text{ if and only if } N(x) \subseteq N(y) \cup \{y\}.$$

We call a set of vertices which are pairwise incomparable with respect to this relation a *vicinal antichain*, and a set of vertices which are pairwise comparable a *vicinal chain*.

The *Dilworth number* of a graph is the maximum size of a vicinal antichain in the graph, or equivalently, the minimum size of a partition of its vertices into vicinal chains.

Split graphs of Dilworth number 1 are precisely *threshold graphs* [Chvátal and Hammer, 1977]. In other words, a graph is threshold if and only if its vertices form a vicinal chain. Let us emphasize that threshold graphs are split graphs, i.e. the vertices of a threshold graph can be partitioned into a clique and an independent set. Moreover, it is not difficult to see that

- (0) if in a split graph with a clique  $C$  and an independent set  $I$  the vertices of  $C$  or the vertices of  $I$  form a vicinal chain, then the set of all vertices of the graph form a vicinal chain.

Universal threshold graphs have been constructed in [Hammer and Kelmans, 1994]. An  $n$ -universal graph in this class contains  $2n$  vertices of which  $n$  vertices form a clique  $C = (c_1, \dots, c_n)$ ,  $n$  vertices form an independent set  $I = (i_1, \dots, i_n)$  and for each  $j$ ,  $N(i_j) = \{c_1, \dots, c_j\}$ . An example of an  $n$ -universal threshold graph is represented in Figure 7.6

Split graphs of Dilworth number at most 2 are precisely split permutation graphs [Benzaken et al., 1985]. In other words, taking into account (0), we can say that a split graph  $G = (C, I, E)$  with a clique  $C$  and an independent set  $I$  is a permutation graph if and only if

- (1) the vertices of  $C$  can be partitioned into at most two sets  $C^1$  and  $C^2$  so that both  $G[C^1 \cup I]$  and  $G[C^2 \cup I]$  are threshold graphs and

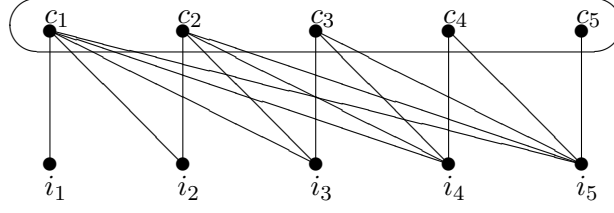


Figure 7.6: Universal threshold graph for  $n = 5$  (the oval represents a clique).

- (2) the vertices of  $I$  can be partitioned into at most two sets  $I^1$  and  $I^2$  so that both  $G[C \cup I^1]$  and  $G[C \cup I^2]$  are threshold graphs.

We say that a split permutation graph  $G = (C, I, E)$  is *symmetric* if it admits a partition  $C = C^1 \cup C^2$  and  $I = I^1 \cup I^2$  such that

- (3) both  $G[C^1 \cup I^1]$  and  $G[C^2 \cup I^2]$  are universal threshold graphs.

If we replaced cliques by independent sets in the definitions of threshold and split permutation graphs we would obtain chain and bichain graphs. More formally:

- A chain graph is a bipartite graph such that vertices in each color class form a chain under neighbourhood inclusion.
- A bichain graph is a bipartite graph such that vertices in each color class can be partitioned into at most 2 sets, each forming a chain under neighbourhood inclusion.

It is known that replacing a clique by an independent set in the  $n$ -universal threshold graph we get an  $n$ -universal chain graph. By inspecting definitions, it is also not hard to see that any  $n$ -universal split permutation graph can be transformed to an  $n$ -universal bichain graph by removing edges from the clique. Therefore, in the rest of this chapter we will only work on constructing an  $n$ -universal split permutation graph. We begin with the following lemma.

**Lemma 7.1.1.** *Every split permutation graph with  $n$  vertices is contained in a symmetric split permutation graph with  $2n$  vertices as an induced subgraph.*

*Proof.* Let  $G = (C, I, E)$  be a split permutation graph with  $n$  vertices and with a partition  $C = C^1 \cup C^2$  and  $I = I^1 \cup I^2$  satisfying (1) and (2). For  $j = 1, 2$ , we denote  $n_j = |C^j| + |I^j|$ . By adding new vertices to the graph  $G[C^1 \cup I^1]$  we can extend it to a universal threshold graph with  $2n_1$  vertices  $C_*^1 = \{c_1^1, c_2^1, \dots, c_{n_1}^1\}$ ,  $I_*^1 = \{i_1^1, i_2^1, \dots, i_{n_1}^1\}$  so that  $N(i_j^1) = \{c_1^1, \dots, c_j^1\}$  for each  $j$ . Similarly, by adding new vertices to the graph  $G[C^2 \cup I^2]$  we can extend it to a universal threshold

graph with  $2n_2$  vertices  $C_*^2 = \{c_1^2, c_2^2, \dots, c_{n_2}^2\}$ ,  $I_*^2 = \{i_1^2, i_2^2, \dots, i_{n_2}^2\}$  so that  $N(i_j^2) = \{c_1^2, \dots, c_j^2\}$  for each  $j$ .

In the graph obtained in this way, condition (3) is satisfied but conditions (1) and (2) are not. To make the new graph a split permutation graph we add some edges between  $I_*^1$  and  $C_*^2$  (to make both of them vicinal chains) and between  $I_*^2$  and  $C_*^1$  (also to make these sets vicinal chains). For  $I_*^1$  and  $C_*^2$  this can be done as follows: whenever there is an edge  $i_k^1 c_j^2$  between  $i_k^1 \in I_*^1$  and  $c_j^2 \in C_*^2$ , we add all missing edges between  $i_k^1$  and any vertex  $c_s^2 \in C_*^2$  with  $s < j$ , and between  $c_j^2$  and any vertex  $i_s^1 \in I_*^1$  with  $s > k$ . For  $I_*^2$  and  $C_*^1$ , the addition of edges can be done by analogy. Since in the original graph the old vertices of each of  $I_*^1$ ,  $C_*^2$ ,  $I_*^2$  and  $C_*^1$  form a vicinal chain, no edge has been added between any two old vertices. Therefore, the resulting graph contains  $G$  as an induced subgraph.  $\square$

Now we establish a correspondence between symmetric split permutation graphs and 321-avoiding permutations. More precisely, we deal with *labelled* 321-avoiding permutations and with *labelled* symmetric split permutation graphs.

By a labelled 321-avoiding permutation we mean a permutation each element of which is assigned one of the two labels 1 or 2 so that the set of 1-labelled elements forms an increasing sequence and the set of 2-labelled elements forms an increasing sequence. Notice that each 321-avoiding permutation  $\pi$  admits at least two possible labellings, while the total number of labellings is  $2^c$ , where  $c$  is the number of connected components of the permutation graph of  $\pi$ . We denote the set of 1-labelled elements of a permutation  $\pi$  by  $\pi^{(1)}$  and the set of 2-labelled elements by  $\pi^{(2)}$ . In other words, we assume that a permutation is labelled if it is given *together* with a partition  $(\pi^{(1)}, \pi^{(2)})$  of its elements into two increasing sequences.

By a labelled symmetric split permutation graph we mean a graph given *together* with a partition of its vertices into four sets  $C^1, C^2, I^1, I^2$  satisfying (1), (2) and (3).

We also adapt the notion of containment relation (pattern containment and induced subgraph containment) to the case of labelled objects (permutations and graphs) in a natural way, i.e. elements of one object belonging to the same set in the given partition must be mapped to elements belonging to the corresponding set in the partition of the other object.

**Lemma 7.1.2.** *There is a one-to-one correspondence  $\phi$  between labelled 321-avoiding permutations and labelled symmetric split permutation graphs. Moreover, a 321-avoiding permutation  $\pi_1$  contains a 321-avoiding permutation  $\pi_2$  as a pattern if and only if  $\phi(\pi_1)$  contains  $\phi(\pi_2)$  as an induced subgraph.*

*Proof.* Suppose we have a labelled 321-avoiding permutation  $\pi = p_1 \dots p_n$  together with a partition  $(\pi^{(1)}, \pi^{(2)})$ . We associate a graph  $G$  with  $2n$  vertices  $C = \{c_1, \dots, c_n\}$  and  $I = \{i_1, \dots, i_n\}$ , where

- $C = C^1 \cup C^2$  with  $C^1 = \{c_{p_j} : p_j \in \pi^{(1)}\}$  and  $C^2 = \{c_{p_j} : p_j \in \pi^{(2)}\}$ ,
- $I = I^1 \cup I^2$  with  $I^1 = \{i_{p_j} : p_j \in \pi^{(1)}\}$  and  $I^2 = \{i_{p_j} : p_j \in \pi^{(2)}\}$ .

We define the set  $C$  to be a clique and the set  $I$  an independent set, i.e.  $G$  is a split graph. The set of edges between  $C$  and  $I$  is defined by describing the neighbourhoods of the vertices of  $I$  as follows:

- If  $p_j \in \pi^{(1)}$ , then  $N(i_{p_j}) = \{c_1, c_2, \dots, c_{p_j}\}$ ,
- if  $p_j \in \pi^{(2)}$ , then  $N(i_{p_j}) = \{c_{p_1}, c_{p_2}, \dots, c_{p_j}\}$ .

By construction,  $I^1$  and  $I^2$  form vicinal chains. Also, from the description it follows that

- for any  $p_j, p_k \in \pi^{(1)}$  with  $p_j < p_k$  we have  $N(c_{p_j}) \supset N(c_{p_k})$  and therefore  $C^1$  forms a vicinal chain,
- for any  $p_j, p_k \in \pi^{(2)}$  with  $p_j < p_k$  we have  $N(c_{p_j}) \supset N(c_{p_k})$  and therefore  $C^2$  forms a vicinal chain.

Therefore,  $G$  is a split permutation graph. Also, it is not difficult to see that

- $|C^1| = |I^1|$  with  $N(i_{p_j}) \cap C^1 = \{c_{p_1}, \dots, c_{p_j}\} \cap \pi^{(1)}$  for each vertex  $i_{p_j} \in I^1$ ,
- $|C^2| = |I^2|$  with  $N(i_{p_j}) \cap C^2 = \{c_{p_1}, \dots, c_{p_j}\} \cap \pi^{(2)}$  for each vertex  $i_{p_j} \in I^2$ .

Therefore,  $G$  is a symmetric split permutation graph. The sets  $C^1, C^2, I^1, I^2$  define a labelling of this graph.

The above procedure defines a mapping  $\phi$  from the set of labelled 321-avoiding permutations of length  $n$  to the set of labelled symmetric split permutation graphs with  $2n$  vertices. It is not difficult to see that different permutations are mapped to different graphs and hence the mapping is injective. Now let us show that this mapping is surjective.

Let  $G = (C, I, E)$  be a symmetric split permutation graph with  $2n$  vertices given together with a partition  $C = C^1 \cup C^2$  and  $I = I^1 \cup I^2$  satisfying (1), (2) and (3). Since  $G$  is symmetric, for any two vertices  $i, j \in I^1$  we have  $|N(i)| \neq |N(j)|$ . Similarly, for any two vertices  $i, j \in I^2$  we have  $|N(i)| \neq |N(j)|$ . Moreover, the degree of each vertex of  $I = I^1 \cup I^2$  is a number between 1 and  $n$ . Now, define

$\pi^{(1)} := \{|N(i)| : i \in I^1\}$  and  $\pi^{(2)} := \{1, 2, \dots, n\} \setminus \pi^{(1)}$  and let  $\pi$  be the permutation obtained by placing the elements of  $\pi^{(2)}$  in the positions  $\{|N(i)| : i \in I^2\}$  in the increasing order, and by placing the elements of  $\pi^{(1)}$  in the remaining positions also increasingly. It is not hard to see that the labelled permutation  $\pi$  with partition  $(\pi^{(1)}, \pi^{(2)})$  is mapped to the graph  $G$ , and thus,  $\phi$  is a surjection.

In conclusion,  $\phi$  is a bijection between labelled 321-avoiding permutations of length  $n$  and labelled symmetric split permutation graphs with  $2n$  vertices. Moreover, this bijection not only maps a permutation  $\pi$  to a graph  $G = (C, I, E)$ , it also maps bijectively the elements of  $\pi$  to  $C$  and to  $I$  as described in the above procedure. This mapping from the elements of the permutation to the vertices of the graph proves the second part of the lemma.  $\square$

**Theorem 7.1.2.** *There is a split permutation graph with  $4n^3$  vertices containing all split permutation graphs with  $n$  vertices as induced subgraphs.*

*Proof.* First, let us define the notion of concatenation of two permutations of the same length. Given two permutations  $\pi_1$  and  $\pi_2$  of length  $n$ , the concatenation  $\pi_1\pi_2$  is the permutation of length  $2n$  obtained by placing the elements of  $\pi_2$  to the right of  $\pi_1$  and by increasing every element of  $\pi_2$  by  $n$ . Also, by  $\pi^n$  we denote the concatenation of  $n$  copies of  $\pi$ .

Let  $\rho_n$  be the  $n$ -universal 321-avoiding permutation of Theorem 7.1.1. Since the permutation graph of  $\rho_n$  is connected (see Figure 7.1), there are exactly two ways to label the elements of  $\rho_n$ . We denote by  $\rho_{n,1}$  and  $\rho_{n,2}$  the two labellings of  $\rho_n$ . Then the concatenation  $\rho_{n,1}\rho_{n,2}$  is universal for labelled 321-avoiding permutations whose permutation graphs are connected. Since every graph with  $n$  vertices contains at most  $n$  connected components, the permutation  $u_n := (\rho_{n,1}\rho_{n,2})^n$  is universal for the set of all labelled 321-avoiding permutations of length  $n$ . Clearly,  $u_n$  is 321-avoiding and hence the graph  $\phi(u_n)$  is a symmetric split permutation graph.

Since  $u_n$  contains all labelled 321-avoiding permutations of length  $n$ , by Lemma 7.1.2  $\phi(u_n)$  contains all labelled symmetric split permutation graphs with  $2n$  vertices. Therefore, by Lemma 7.1.1,  $\phi(u_n)$  is an  $n$ -universal split permutation graph. Since the length of  $\rho_n$  is  $n^2$ , the length of  $u_n$  is  $2n^3$ , and hence the number of vertices of  $\phi(u_n)$ , is  $4n^3$ .  $\square$

### 7.1.3 Z-grid

In this section we will provide with an explicit construction of an  $n$ -universal split permutation and an  $n$ -universal bichain graphs. To this end we introduce a specific

bichain graph  $Z_{n,k}$  consisting of  $n \times k$  vertices organized in a grid-like structure with  $n$  rows and  $k$  columns. The graph  $Z_{6,7}$  is a graph represented in Figure 7.7 together with so called "diagonal" edges connecting every even column  $j$  to every odd column  $j' \geq j + 3$  (these edges are not shown for clarity of the picture).

We call any graph of the form  $Z_{n,k}$  a  $Z$ -grid. More formally:

**Definition 7.1.1.** Define  $Z_{n,k}$  to be a graph with the vertex set  $V(Z_{n,k}) = \{z_{i,j} : 1 \leq i \leq n, 1 \leq j \leq k\}$  with  $z_{i,j}z_{i',j'}$  an edge if and only if

- $j$  is even,  $j'$  is odd and  $j' \geq j + 3$
- $j$  is odd and  $j' = j + 1$  and  $i > i'$
- $j$  is even and  $j' = j + 1$  and  $i \leq i'$ .

By adding edges between vertices in even columns we obtain a corresponding split permutation graph:

**Definition 7.1.2.** Define  $Z_{n,k}^*$  to be a graph obtained from  $Z_{n,k}$  by adding edges between  $z_{i,j}z_{i',j'}$  if

- $j$  and  $j'$  are even.

In this section we will prove that these graphs are universal:

**Theorem 7.1.3.** *For each  $n$ , the graph  $Z_{n,n+1}$  is an  $n$ -universal bichain graph and the graph  $Z_{n,n+1}^*$  is an  $n$ -universal split permutation graph.*

**Definition 7.1.3.** Let the  $\rho_n$  be a universal 321-avoiding permutation defined in Chapter 7.1.1. We arrange the entries of the permutation into the  $n \times n$  table  $\{A(x, y) : 1 \leq x, y \leq n\}$  according to the picture 7.1, i.e.:

- $A(x, 1) = x$ ,
- $A(x, y) = (2ny - 3n) + 2x$  for  $y$  even and  $y < n$ ,
- $A(x, y) = (2ny - 3n) + 2x - 1$  for  $y$  odd,
- $A(x, n) = n^2 - n + x$  if  $n$  is even.

Then we define two graphs  $G_n^1$  and  $G_n^2$  on the vertex set  $\{i_{A(x,y)}, c_{A(x,y)} : 1 \leq x, y \leq n\}$  so that

- the vertices  $\{i_{A(x,y)} : 1 \leq x, y \leq n\}$  form an independent set and the vertices  $\{c_{A(x,y)} : 1 \leq x, y \leq n\}$  form a clique;



- If  $y$  is odd (resp. even) then  $i_{A(x,y)}c_{A(x',y')}$  is an edge in  $G_n^1$  (resp.  $G_n^2$ ) if and only if  $A(x,y) \geq A(x',y')$ .
- If  $y$  is even (resp. odd) then  $i_{A(x,y)}c_{A(x',y')}$  is an edge in  $G_n^1$  (resp.  $G_n^2$ ) if and only if  $A(x,y)$  comes before entry  $A(x',y')$  in the permutation  $\rho_n$ , i.e.  $\rho_n^{-1}\{A(x,y)\} \leq \rho_n^{-1}\{A(x',y')\}$

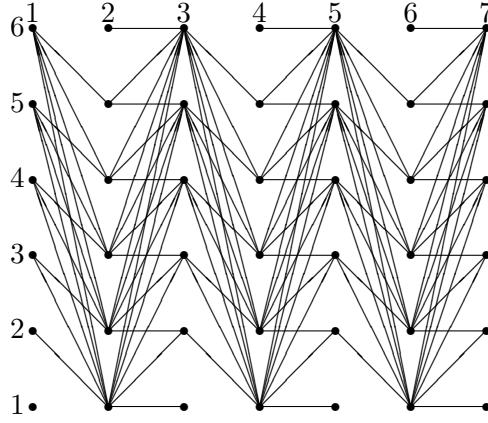


Figure 7.7: Graph  $Z_{6,7}$  without the "diagonal" edges

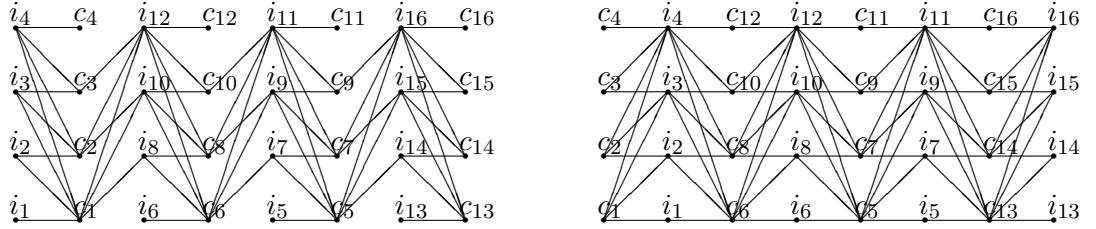


Figure 7.8: Graphs  $G_4^1 = \phi(\rho_4^1)$  (left) and  $G_4^2 = \phi(\rho_4^2)$  (right) without "diagonal" and "clique" edges

Now, it is not hard to see that entries in the even columns and entries in the odd columns of the matrix  $A(x,y)$  both form increasing sequences and therefore corresponds to the partition of permutation  $\rho_n$ . Hence  $G_n^1 = \phi(\rho_n^1)$  and  $G_n^2 = \phi(\rho_n^2)$ .

Notice that if we arranged the entries of  $G_n^1$  and  $G_n^2$  in a grid-like structure according to matrix  $A(x,y)$ , the resulting graph would resemble the  $Z$ -grid. In Figure 7.8 we picture graphs  $\phi(u_4^1)$  and  $\phi(u_4^2)$  with the "diagonal" edges between  $i_k$  and  $c_{k'}$  whenever the column of  $c_{k'}$  is to the left and not adjacent to the column of  $i_{k'}$  and the "clique" edges connecting  $c_k$ 's omitted for clarity. We can check that

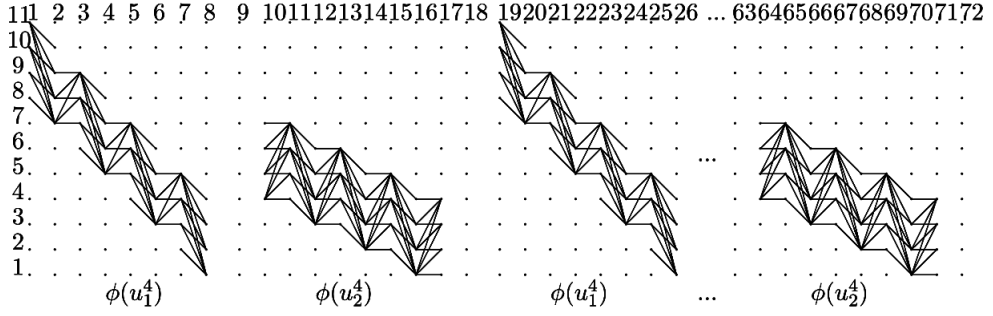


Figure 7.9:  $Z_{11,72}^*$  contains the universal bichain graph  $\phi((\rho_4^1 \rho_4^2)^4)$

the picture on the left (resp. right) is indeed  $G_4^1$  (resp.  $G_4^2$ ), because the edges are obtained by joining the vertices  $i_k$  in the 1st and 5th columns (resp. 4th and 8th columns) to all  $c_{k'}$ 's with  $k' \leq k$ . For example, in the picture on the left  $i_9$  is joined to  $c_5, c_6, \dots, c_9$  and to  $c_1, c_2, c_3, c_4$  by diagonal edges. On the other hand, the vertices  $i_k$  on the 3rd and 7th column (resp. 2nd and 6th column) is joined to all  $c_{k'}$ 's with  $k'$  appearing to the left of corresponding index of  $k$  in the universal permutation  $\rho_4 = 6182(10)3(12)4(13)5(14)7(15)9(16)(11)$ . For example, in the picture on the left,  $i_{14}$  is not joined only to  $c_7, c_9, c_{11}, c_{15}, c_{16}$  as only 7, 9, 11, 15 and 16 appears to the right of 14 in universal permutation  $\rho_4$ .

Now, by direct inspection we can claim that  $G_n^1$  and  $G_n^2$  can be embedded into  $Z$ -grid of size  $Z_{3n-1,4n+2}^*$  and having  $n$  copies of these embeddings we get an embedding of universal graph  $\phi(u_n)$  into  $Z_{3n-1,4n^2+2n}^*$  (recall that  $u_n = (\rho_n^1 \rho_n^2)^n$ ). In Figure 7.9 we show how 4 copies of  $G_4^1$  and  $G_4^2$  and hence  $\phi(u_n)$  can be embedded into  $Z_{11,72}$ . Notice that  $i_k$ 's are embedded into odd columns and  $c_k$ 's are embedded into even columns, therefore the "diagonal" and "clique" edges between columns in  $Z$ -grid coincide with the "diagonal" and "clique" edges between columns of universal bichain graph.

We can generalize this construction to show that a copy of  $G_n^1$  can be found in a sliding strip  $\{z_{x,y} : y \in \{1, 2, \dots, 2n\}, x + y = \{2n + 1, 2n + 2, \dots, 3n\}\}$  and a copy of  $G_n^2$  can be found in a strip with twice smaller slope  $\{z_{x,y} : y \in \{2n + 2, 2n + 3, \dots, 4n + 1\}, \lfloor \frac{j}{2} \rfloor + x \in \{2n + 1, 2n + 2, \dots, 3n\}\}$ . These strips can be shifted to the right  $n$  times by adding  $4n + 2$  to the first coordinate of  $z_{i,j}$  each time. This way we get the embedding of  $\phi((\rho_n^1 \rho_n^2)^n)$  into  $Z_{3n-1,4n^2+2n}^*$ . Notice first that  $i_k$ 's of  $G_n^1$  and  $G_n^2$  are embedded into odd columns and  $c_k$ 's into even columns. Therefore the additional edges between  $i_k$ 's and  $c_{k'}$ 's coincide with the additional edges of  $Z$ -grid.

This proves that  $Z$ -grid is an  $n$ -universal graph, i.e. any  $n$ -vertex split per-

mutation graph can be embedded into grid  $Z_{3n-1,4n^2+2n}^*$ . Now remark that deletion of any row in a  $Z$ -grid leaves us with a  $Z$ -grid with smaller number of rows. Hence, for any embedding of any  $n$ -vertex graph into  $Z$ -grid, we can obtain an embedding into a  $Z$ -grid with at most  $n$  rows by deleting all rows not used in embedding. Hence  $Z_{n,4n^2+2n}^*$  is universal.

The deletion of a column which is neither the first nor the last in a  $Z$ -grid, however, does not leave us with a  $Z$ -grid. Nevertheless, the rest of the graph can be embedded into a  $Z$ -grid with a smaller number of columns. In Figure 7.10 we show how after removing a column of  $Z_{n,m}$ , we can arrange the remaining vertices (parts  $A$  and  $B$  in the picture) inside  $Z_{2n,m-2}$ . Hence, by applying this operation repeatedly for every empty column, we can transform any embedding of  $n$ -vertex graph into  $Z$ -grid into another one for which the vertices occupy consecutive columns. Further, we can assume that the last column and the first or the second is not empty. Thus every  $n$ -vertex split permutation graph can be embedded into a  $Z$ -grid with at most  $n + 1$  columns. If we now delete unused rows as described above, we obtain an embedding into  $Z_{n,n+1}^*$ . Thus  $Z_{n,n+1}^*$  is a universal split permutation graph. By removing the clique, we obtain that  $Z_{n,n+1}$  is a universal bichain graph and hence we are done.

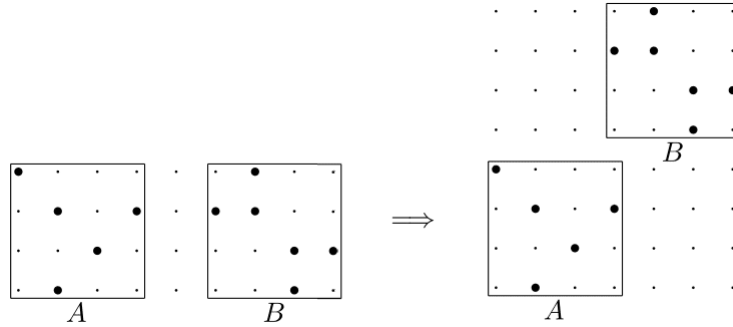


Figure 7.10: Embedding a graph into a  $Z$ -grid with fewer columns

## 7.2 Minimality of classes of bichain and split permutation graphs

### 7.2.1 Bipartite permutation graphs and the $X$ -grid

Lozin [2011] proved that the class of bipartite permutation graphs is a minimal hereditary class of unbounded clique-width. In the proof of minimality, the author

used the construction of universal bipartite permutation graph  $X_{n,n}$  (Figure 7.11) and proved the following theorem:

**Theorem 7.2.1.** *For every  $n$ , the clique-width of  $X_{n,n}$ -free bipartite permutation graphs is bounded by a constant.*

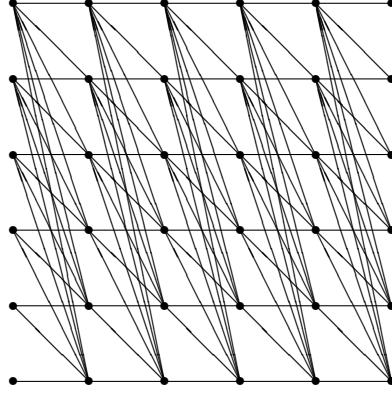


Figure 7.11: Graph  $X_{6,6}$

In this chapter we will prove that bichain graphs and split permutation graphs are minimal classes of unbounded clique-width. We will prove minimality for the class bichain graphs by pivoting bichain graphs into bipartite permutation graphs and then deduce the result for split permutation graphs. As an intermediate step in this process, we use more grid-like graphs. One of them is denoted  $Y_{n,k}$  ( $n$  columns and  $k$  rows) and an example of this graph with  $n = 7$  and  $k = 5$  is shown in Figure 7.12 (left). By adding to  $Y_{n,k}$  an extra line at the bottom as shown in Figure 7.12 (right) we obtain a grid which we denote  $Y_{n,k}^+$ . The example shown in Figure 7.12 (right) represents the graph  $Y_{n,k}^+$  with  $n = 7$  and  $k = 5$  (note that  $Y_{n,k}^+$  contains  $k + 1$  rows). We will refer to graphs of the form  $Y_{n,k}$  and  $Y_{n,k}^+$  as  $Y$ -grids and  $Y^+$ -grids, respectively.

It is not difficult to see that the  $Y$ -grid is simply a modification of the  $X$ -grid obtained by shifting the vertices within each column so that every horizontal line of the  $X$ -grid turns  $45^\circ$  clockwise. Therefore, any  $Y$ -grid is a bipartite permutation graph and hence can be embedded into an  $X$ -grid. With a bit of care one can also see that a  $Y^+$ -grid is embeddable into an  $X$ -grid as well. Figure 7.13(left) illustrates how  $Y_{3,3}^+$  can be embedded into  $X_{6,6}$ . On the other hand, an  $X$ -grid can be embedded into a  $Y$ -grid and hence into a  $Y^+$ -grid, as exemplified in Figure 7.13(right).

Both examples can be easily generalized to the following statement.

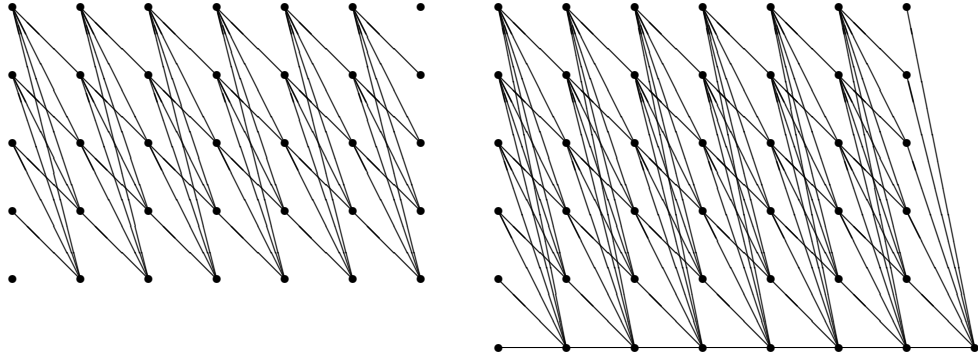


Figure 7.12: Graphs  $Y_{7,5}$  (left) and  $Y_{7,5}^+$  (right)

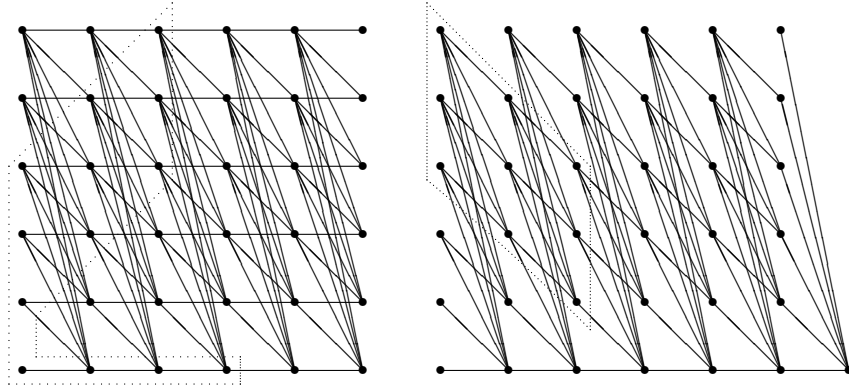


Figure 7.13: The graph  $X_{6,6}$  contains the graph  $Y_{3,3}^+$  (left) and the graph  $Y_{6,5}^+$  contains the graph  $X_{3,3}$  (right)

**Lemma 7.2.1.** *The graph  $X_{2n,2n}$  contains  $Y_{n,n}$  and  $Y_{n,n}^+$  as induced subgraphs, and the graphs  $Y_{2n,2n}$  and  $Y_{2n,2n}^+$  contain  $X_{n,n}$  as an induced subgraph.*

In the proof of the main result we will use the property of  $X$ -grid described by the lemma below. We first start by the definition:

**Definition 7.2.1.** Let  $X_{n,n}$ ,  $X_{m,m}$  be two  $X$ -grids with vertex sets  $\{v_{i,j} | 1 \leq i, j \leq n\}$  and  $\{w_{i,j} | 1 \leq i, j \leq m\}$  respectively, and edges between  $v_{i,j}$  and  $v_{i',j'}$  (and between  $w_{i,j}$  and  $w_{i',j'}$ , respectively) if and only if  $i' = i + 1$  and  $j' \geq j$  or  $i = i' + 1$  and  $j \geq j'$ . Suppose  $m \geq n$  and that there is an embedding  $\phi : X_{n,n} \rightarrow X_{m,m}$  mapping vertices  $v_{i,j}$  injectively to vertices  $w_{\phi_1(i,j), \phi_2(i,j)}$ , so that edges are mapped to corresponding edges. We call a mapping  $\phi$  forward (resp. reverse) monotonous if and only if  $\phi_1(i, j) = i + \phi_1(1, 1) - 1$ ,  $\phi_2(i, j + 1) > \phi_2(i, j)$  and  $\phi_2(i + 1, j) \geq \phi_2(i, j)$  ( resp.  $\phi_1(i, j) = i + \phi_1(1, 1) - 1$ ,  $\phi_2(i, j + 1) < \phi_2(i, j)$  and  $\phi_2(i + 1, j) \leq \phi_2(i, j)$ ). A mapping which is forward or reverse monotonous will be called monotonous.

Now we can state a lemma which will be used in the final step of our main result:

**Lemma 7.2.2.** *Suppose the  $X$ -grid  $X_{n,n}$  contains an  $X$ -grid  $X_{m,m}$  as an induced subgraph under embedding  $\phi : v_{i,j} \rightarrow w_{\phi_1(i,j), \phi_2(i,j)}$ . Then the mapping  $\phi$  restricted to either the set of vertices  $S_1 = \{v_{i,j} | \frac{2n}{3} \leq i \leq n, 1 \leq j \leq \frac{n}{3}\}$  or to the set of vertices  $S_2 = \{v_{i,j} | \frac{2n}{3} \leq i \leq n, \frac{2n}{3} \leq j \leq n\}$  is a monotonous embedding.*

*Proof.* Consider a vertex  $v_{i,1}$  in the row  $2 \leq i \leq n-1$  which is mapped to some vertex  $w_{i',j'} = w_{\phi_1(i,1), \phi_2(i,1)}$ . Now,  $v_{i,1}$  has  $n+1$  neighbours,  $N = \{v_{i-1,1}, v_{i+1,1}, v_{i+1,2}, \dots, v_{i+1,n}\}$ . Then, for some  $k \in \{0, 1, 2, \dots, n+1\}$  we will have  $k$  neighbours  $\phi(N)$  will lie in  $i' - 1$ 'th column, i. e. the set  $\{w_{i'-1,j_1}, w_{i'-1,j_2}, \dots, w_{i'-1,j_k}\}$  and other  $n+1-k$  neighbours will lie in  $i' + 1$ 'st column, i. e.  $w_{i'+1,j_{k+1}}, w_{i'+1,j_{k+2}}, \dots, w_{i'+1,j_{n+1}}$  in column  $i' + 1$ . Moreover, we can assume that  $j_1 < j_2 < \dots < j_k \leq j' \leq j_{k+1} < j_{k+2} < \dots < j_n$ .

Now, we define sets of vertices  $C_{l,p} = \{w_{i,j} : i = l, j \in (j_p, j_{p+1}]\}$ , for  $p = 1, 2, \dots, n$  and  $C_{l,n+1} = \{w_{i,j} : i = l, j \in (j_{n+1}, m) \cup i = l - 2, j \in [1, j_1]\}$ . And define intervals  $C_{l,p}^r = \{w_{i,j} : i = l, j \in [j_{p-1}, j_p)\}$ , for  $p = 2, 3, \dots, n+1$  and  $C_{l,1} = \{w_{i,j} : i = l, j \in j < j_1 \cup i = l + 2, j \in [j_n, m]\}$  Now, we look where  $v_{i,2}$  can be embedded. Notice, that  $v_{i,1}$  and  $v_{i,2}$  have  $n$  common neighbours and one private neighbour. The only possible positions for this to happen are

- $\phi(v_{i,2}) \in C_{i',k+1}$ .
- $\phi(v_{i,2}) \in C_{i',k}^r$

Now, as  $v_{i,2}$  and  $v_{i,3}$  differ in just one neighbour, we can continue this procedure revealing where the whole column is embedded. The result would be:

- $\phi(v_{i,j} \in C_{i',k+j-1}$  for  $j = 1, 2, \dots, n+2-k$  and  $\phi(v_{i,j} \in C_{i'-2,k+j-n-2}$  for  $j = (n+3-k), (n+4-k), \dots, n\}$
- $\phi(v_{i,j} \in C_{i',k+1-j}^r$  for  $j = 1, 2, \dots, k+1$  and  $\phi(v_{i,j} \in C_{i'+2,k+n+2-j}^r$  for  $j = (k+2), (k+3), \dots, n\}$ .

Now we notice that in both cases we have that  $vi-1, 1$  has  $k$  or  $k-1$  neighbours and the column  $i-1$  must be embedded according to the previous rule. Hence we apply this rule repeatedly finding where columns  $n, n-1, \dots, 2n/3$  are embedded. Then looking at the initial  $k$ , we can the last  $k$  will be in the interval  $[k, k-n/3]$  and hence depending on whether  $k \geq 2n/3$  or  $k < 2n/3$  we have that intervals  $[1, n/3]$  or  $[2n/3, n]$  will not contain any values of  $k$ . These intervals give the corresponding regions where mapping is monotonous. This proves the lemma.  $\square$

### 7.2.2 Pivoting bichain graphs into bipartite permutation graphs

Let  $ab$  be an edge in a bipartite graph,  $A = N(a) - \{b\}$  and  $B = N(b) - \{a\}$ . The operation of pivoting consists in complementing the edges between  $A$  and  $B$ . i.e. replacing every edge  $xy$  ( $x \in A$  and  $y \in B$ ) with a non-edge and vice versa.

The goal of this section is to show that a  $Z$ -grid can be transformed by a sequence of pivoting operations into a bipartite permutation graph. Since the pivoting operation is very sensible, we apply it to a particular form of the  $Z$ -grid. First, we restrict ourselves to graphs of the form  $Z_{n,k}$  only with odd values of  $n$ . Second, we extend  $Z_{n,k}$  in a specific way by adding to it an extra bottom line and denote the resulting graph by  $Z_{n,k}^+$ . Formally speaking, the graph  $Z_{n,k}^+$  can be obtained from the graph  $Z_{n+2,k+1}$  by deleting the bottom-left vertex and all vertices of the last two columns except the bottom vertices. An example of the graph  $Z_{n,k}^+$  with  $n = 7$  and  $k = 5$  is shown on the left of Figure 7.14. The bottom line is labelled by 0. We will call the bottom line of the graph  $Z_{n,k}^+$  the pivoting line, which is explained by the following lemma.

**Lemma 7.2.3.** *The sequence of pivoting operations applied to the bottom edges of the graph  $Z_{n,k}^+$  in the order from left to right starting from the second edge transforms  $Z_{n,k}^+$  into the graph  $Y_{n,k}^+$ .*

*Proof.* The result follows by direct inspection. The only peculiarity of this transformation is that after the pivoting every vertex of the bottom line in an odd column  $i$  moves to column  $i-2$ , as illustrated in Figure 7.14.  $\square$

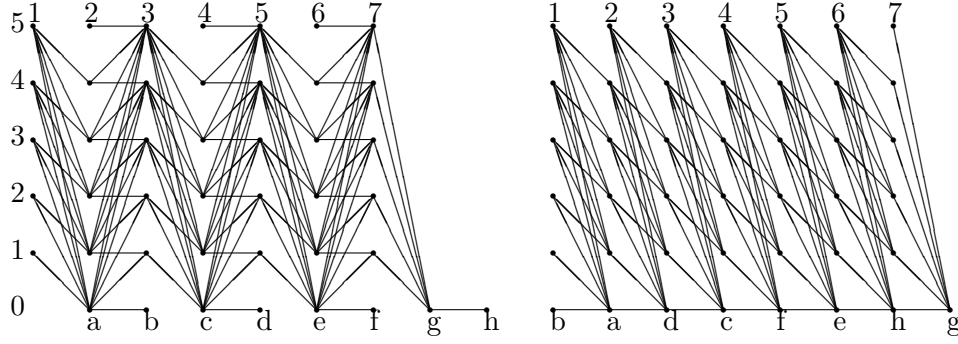


Figure 7.14: The graph  $Z_{7,5}^+$  with "diagonal" edges omitted (left) transforms into the graph  $Y_{7,5}^+$  (right) by pivoting on the edges  $cd$ ,  $ef$ ,  $gh$ .

### 7.2.3 Proof of minimality

We will use the following result proved in [Kamiński et al., 2009]:

**Lemma 7.2.4.** *For a class of graphs  $X$  and a non-negative integer  $k$ , denote by  $X^{(k)}$  a class of graphs obtained from class in  $X$  by at most  $k$  subgraph complementations. Then the clique-width of graphs in  $X^{(k)}$  is bounded if and only if the clique-width of graphs in  $X$  is bounded.*

**Theorem 7.2.2.** *The classes of bichain graphs and of split permutation graphs are a minimal hereditary classes of unbounded clique-width.*

*Proof.* It was proved in [Korpelainen et al., 2014] that the class of split permutation graphs has unbounded clique-width. By Lemma 7.2.4 it follows that the class of bichain graphs has unbounded clique-width as well.

Now, let  $\mathcal{A}$  be a proper hereditary subclass of bichain graphs, i.e. a subclass obtained by forbidding at least one bichain graph. According to Lemma 7.1.3, every graph  $G \in \mathcal{A}$  can be embedded into a  $Z$ -grid  $Z_{N,N}$ . We extend  $Z_{N,N}$  to  $Z_{N,N}^+$  by adding to it a pivoting line. We also add the pivoting line to  $G$  and denote the resulting graph by  $G_+$ .

By pivoting on the edges of the pivoting line, we transform  $G_+$  into a graph  $G_+^*$ , and by deleting the pivoting line from  $G_+^*$  we obtain a graph which we denote by  $G^*$ . According to Lemma 7.2.1,  $G_+^*$  (and hence  $G^*$ ) is a bipartite permutation graph.

For an arbitrary bipartite permutation graph  $H$ , let  $x(H)$  be the maximum  $n$  (the  $x$ -number) such that  $H$  contains  $X_{n,n}$  as an induced subgraph, and for an arbitrary bichain graph  $H$  we denote by  $z(H)$  the maximum  $n$  (the  $z$ -number) such that  $H$  contains  $Z_{n,n}$  as an induced subgraph. Denote



$$\mathcal{A}_+ = \{G_+ : G \in \mathcal{A}\}$$

$$\mathcal{A}_+^* = \{G_+^* : G_+ \in \mathcal{A}_+\}$$

$$\mathcal{A}^* = \{G^* : G_+^* \in \mathcal{A}_+^*\}$$

Assume the  $x$ -number is unbounded for graphs in  $\mathcal{A}^*$ . Then by Lemmas 7.2.3 and 7.2.1 the  $z$ -number is unbounded for graphs in  $\mathcal{A}$ . But then by Lemma 7.1.3  $\mathcal{A}$  must contain all bichain graphs, contradicting our assumption. This contradiction shows that the  $x$ -number is bounded by a constant, say  $k$ , for graphs in  $\mathcal{A}^*$ , i.e. these graphs are  $X_{k,k}$ -free. Therefore, graphs in  $\mathcal{A}_+^*$  are  $H_{k+1,k+1}$ -free, since by adding one line of the grid we can increase the  $x$ -number by at most one. By Theorem 7.2.1 this implies that graphs in  $\mathcal{A}_+^*$  have bounded clique-width. Therefore, they also have bounded rank-width. Since pivoting does not change rank-width, graphs in  $\mathcal{A}_+$  have bounded rank, and hence, clique-width. As a result, graphs in  $\mathcal{A}$  have bounded clique-width.

This proves that the class of bichain graphs is minimal class of unbounded clique-width. By Lemma 7.2.4 the minimality follows for the class of split permutation graphs and hence we are done.  $\square$

### 7.3 Canonical antichain

In this section we will show that the class of bichain graphs contains a canonical antichain with respect to the labelled induced subgraph relation and then we will remark on how this argument can be modified to yield a canonical antichain for split permutation graphs.

The canonical antichain for the class bichain graphs consist of graphs  $S_k$  defined as follows:

**Definition 7.3.1.** Define the graph  $S_k$  to be a bipartite graph on vertex set  $\{x_i, y_i : 1 \leq i \leq k\}$  with  $x_i y_j \in E(S_k)$  if and only if  $j = i$  or  $j \geq i + 2$ .

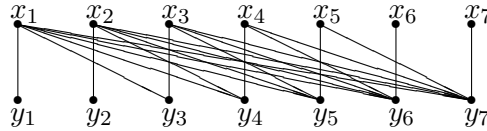


Figure 7.15: The graph  $S_7$

The graph  $S_7$  is pictured in Figure 7.15.

First observe that the graphs  $S_k$  belong to the class of bichain graphs. Indeed, the neighbourhoods of the vertices with even indices and the neighbourhoods of the vertices of odd indices split each part of bipartition into two chains. Also, notice that the only induced  $2K_2$ 's in these graphs are formed by the vertices in two consecutive columns. Hence, any embedding of  $S_k$  into  $S_{k'}$  must map the vertices with consecutive indices to the vertices of consecutive indices. By labelling the vertices in the first and the last column of  $S_k$  differently from the vertices in other columns, we deduce that graphs  $S_k$  form an antichain with respect to labelled induced subgraph relation.

In this section we will prove that, in fact, these graphs constitute a canonical antichain. In other words, we will prove that every proper subclass of bichain graphs, which contains only finitely many graphs  $S_k$ , is well-quasi-ordered by the labelled induced subgraph relation.

Let us recall the notion of letter graphs discussed in Section 3.2:

**Definition 7.3.2.** A  $k$ -letter graph  $G$  is a graph defined by a finite word  $x_1x_2 \dots x_n$  on alphabet  $X$  of size  $k$  together with a subset  $S \subseteq X^2$  such that:

- $V(G) = \{x_1, x_2, \dots, x_n\}$
- $E(G) = \{x_i x_j : i \leq j \text{ and } (x_i, x_j) \in S\}$

The importance of this notion is due to the following theorem proved in [Petkovšek, 2002]:

**Theorem 7.3.1.** *For any fixed  $k$ , the class of  $k$ -letter graphs is well-quasi-ordered by the labelled induced subgraph relation.*

Let us start by proving that bichain graphs which can be embedded into  $k$  consecutive columns of a  $Z$ -grid are  $k$ -letter graphs.

**Lemma 7.3.1.** *For any  $n, k \in \mathbb{N}$ ,  $Z_{n,k}$  is a  $k$ -letter graph.*

*Proof.* Let  $A = \{a_1, a_2, \dots, a_k\}$ , let  $R_1 = \{a_1 a_2, a_3 a_4, \dots, a_{2\lfloor \frac{k}{2} \rfloor - 1} a_{2\lfloor \frac{k}{2} \rfloor}\}$ ,  $R_2 = \{a_3 a_2, a_5 a_4, \dots, a_{2\lceil \frac{k}{2} \rceil - 1} a_{2\lceil \frac{k}{2} \rceil - 2}\}$ ,  $D = \{a_{2i} a_{2j+1}, a_{2j+1} a_{2i} : 1 \leq i < j \leq \lceil \frac{k}{2} \rceil - 1\}$  and let the word be  $w = (a_k a_{k-1} \dots a_1)^n$ . Then it is not hard to see that the  $k$ -letter graph associated with the word  $w$  is isomorphic to  $Z_{n,k}$  with the letters of the word  $w$  corresponding to the vertices of  $Z$ -grid as pictured in Figure 7.16.  $\square$

In this section we will also use canonical decomposition of bipartite graphs discussed in Section 3.1.5 and the Theorem 3.1.3 stating that to prove well-quasi-ordering for a class of bipartite graphs it is enough to prove it for the set of canonically prime graphs in the class.

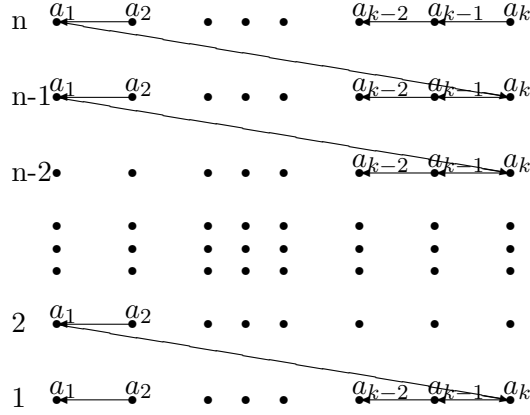


Figure 7.16: Reading  $Z_{n,k}$  as a  $k$ -letter graph

In the rest of the proof we will show that each canonically prime graph not containing  $S_k$  as an induced subgraph must be embeddable into  $Z_{n,2k+1}$  for some  $n$ . Hence, by Lemma 7.3.1 and Theorem 7.3.1 canonically prime graphs in the class  $Free(S_k)$  are well-quasi-ordered by the labelled induced subgraph relation. By Theorem 3.1.3 this implies that the class  $Free(S_k)$  is well-quasi-ordered by labelled induced subgraph relation. This would imply that any hereditary subclass of bichain graphs which contains finitely many subgraphs from the antichain  $S_k$  must be well-quasi-ordered, i. e.  $S_k$  is a canonical antichain.

**Definition 7.3.3.** An embedding of a bichain graph  $G$  into a  $Z$ -grid is a function  $e : V(G) \rightarrow \mathbb{N}^2$  such that for any two vertices  $v, w \in V(G)$  we have  $vw \in E(G)$  if and only if  $e(v)e(w)$  is an edge in some  $Z_{n,k}$  (with  $n$  and  $k$  large enough for  $V(Z_{n,k})$  to contain  $e(V(G))$ ).

First of all, notice that each bichain graph has an embedding into a  $Z$ -grid by Lemma 7.1.3. Also, it's not hard to see that we can always recover a graph from any of its embeddings.

**Definition 7.3.4.** Let  $z$  be a function which maps a subset  $S \subset \mathbb{N}^2$  to the graph  $z(S)$  with vertex set  $S$  and  $(i, j)(i', j') \in z(S)$  if and only if  $(i, j)(i', j') \in Z_{n,k}$  (with  $n$  and  $k$  large enough for  $V(Z_{n,k})$  to contain  $S$ ).

So given a bichain graph  $G$  and any of its embeddings  $e(G)$ , we see that  $z(e(G))$  is isomorphic to  $G$ . Hence, in what follows we will analyze the structure of embeddings of  $S_k$ -free canonically prime bichain graphs. To start with, let us find out some subsets  $S \in \mathbb{N}^2$  such that  $f(S) = S_k$ .

**Definition 7.3.5.** Let  $s_1, s_2, \dots, s_r \in \mathbb{N}$  with  $s_i \geq 5$  for all  $i$ . Let  $S = \{(x_{ij}, y_{ij}) \in \mathbb{N}^2 : 1 \leq i \leq r, 1 \leq j \leq s_i\}$  be a set of points in  $\mathbb{N}^2$  satisfying the following:

- $x_{i(j+1)} = x_{ij} + 1$ ;
- $y_{i(j+1)} < y_{ij}$ ;
- $x_{(i+1)1} = x_{is_i} - 3$ ;
- $y_{(i+1)j} \geq y_{i(s_i-5+j)}$  for  $j = 1, 2, 3, 4, 5$ .

We will call such sets good of size  $|S|$ .

The first and the second condition says that for each  $i$ , the set of points  $\{(x_{ij}, y_{ij}) : 1 \leq j \leq s_i\}$  constitutes a decreasing sequence in the consecutive columns from  $x_{i1}$  to  $x_{is_i}$ . Further, the third and fourth conditions explain that the  $i+1$ 'th and  $i$ 'th sequence share exactly 4 columns and that the  $y$  coordinates of  $i+1$ 'th sequence are above the  $y$  coordinates of the  $i$ 'th sequence in the columns shared. Moreover, last condition says more than that, it says that the entries of  $i$ 'th sequence cannot be replaced by some entries of  $i+1$ 'th sequence to make a decreasing sequence in the consecutive columns. The motivation of this definition is due to the following lemma:

**Lemma 7.3.2.** *Let  $S$  be as in definition above with  $x_{11}$  odd and  $|S| = 2k$ . Then  $f(S)$  is isomorphic to  $S_k$ .*

*Proof.* By inspection. □

The following lemma completes the proof:

**Lemma 7.3.3.** *Let  $G$  be a canonically prime bichain graph and  $e(G)$  be an embedding of  $G$  which contains two vertices  $(x, y)$  and  $(x', y')$  such that  $|x' - x| \geq 2k + 1$ , then it contains a good set of size at least  $2k+1$ .*

Indeed, Lemmas 7.3.2 and 7.3.3 imply that any  $S_k$ -free canonically prime bichain graph has an embedding occupying at most  $2k$  consecutive columns of  $Z$ -grid and hence can be embedded into  $Z_{2k+1, n}$  for some  $n$ . By the previous discussion the conclusion follows.

To prove Lemma 7.3.3 we need the following technical definitions.

**Definition 7.3.6.** Let  $S \in \mathbb{N}^2$ , and let  $(i, j), (i', j') \in S$ , then  $(i', j')$  is called a neighbour of  $(i, j)$  if  $i' = i + 1$  and  $j' < j$  and for every  $j''$  such that  $(i + 1, j'') \in S$  and  $j'' < j$  we have  $j'' \leq j'$ . We call the two points  $(i, j)$  and  $(i', j')$  a neighbouring pair in  $S$ .

**Definition 7.3.7.** Let  $S \in \mathbb{N}^2$ . Then we construct a graph  $S^{\mathcal{N}}$  which has vertex set  $S$  and an edge between two vertices if and only if they are neighbouring pair in  $S$ . We call the graph  $S^{\mathcal{N}}$  a neighbour graph of  $S$ .

**Definition 7.3.8.** For a set  $S \in \mathbb{N}^2$ , denote by  $l_S$  (resp.  $r_S$ ) the number of the leftmost (resp. rightmost) column which has a point in  $S$ , i.e.  $l_S = \min\{x : (x, y) \in S\}$  (resp.  $r_S = \max\{x : (x, y) \in S\}$ ). Also, denote by  $lh_S$  the highest point in the leftmost column belonging to  $S$ , i.e.  $lh_S = (l_S, h_S)$  where  $h_S = \max\{y : (l_S, y) \in S\}$ .

Now we proceed to the proof of Lemma 7.3.3:

*Proof.* Take a canonically prime bichain graph  $G$  and an embedding into a  $Z$ -grid  $e$ . We set  $B_0 = e(G)$ ,  $A_0 = \emptyset$  and recursively for  $i \geq 1$  we define the following subsets of  $e(G)$ :

- If  $B_{i-1} = \emptyset$  then stop, otherwise define the following three sets:
- Let  $C_i$  be the component of  $B_{i-1}^{\mathcal{N}}$  containing  $lh_{B_{i-1}}$
- Let  $A_i$  be the union of points in  $A_{i-1}$  and the points of  $e(G)$  in the columns of  $C_i$  which lie 'below  $C_i$ ', i.e. the set  $\{(x, y) \in e(G) : \exists (x', y') \in C_i, x = x', y \leq y'\}$ .
- Let  $B_i$  be the complement of  $A_i$  in  $e(G)$ , i. e.  $B_i = e(G) \setminus A_i$ .

For easier reading let us denote  $l_i = l_{C_i}$  and  $r_i = r_{C_i}$ . Now we will show that  $r_{i+1} \leq r_i$  and  $l_{i+1} \leq r_i - 3$ . Suppose by contradiction  $r_{i+1} < r_i$  and let  $(r_{i+1}, y)$  be any point in  $r_{i+1}$ th column belonging to  $C_{i+1}$ . Then, by definition  $r_{i+1} \geq l_{i+1} > l_i$  and hence the point  $(r_{i+1}, y)$  lies above some point  $(r_{i+1}, y')$  belonging to  $C_i$  with  $y > y'$ . Since  $(r_{i+1}, y')$  has a neighbour in  $e(G)$ , the point  $(r_{i+1}, y)$  must have a neighbour in  $e(G)$  which contradicts the definition of  $r_{i+1}$ . Now, if  $l_{i+1} \geq r_i - 2$ , then it is not hard to see the sets  $A_i$  and  $B_i$  define a skew-join partition of  $G = z(A_i) \odot z(B_1)$  (every entry of even column of  $A_i$  is joined to every entry of with odd column of  $B_i$  and there are no other edges between parts  $A_i$  and  $B_i$  in  $G$ ).

Finally, we observe that each component  $C_i$  contains a decreasing sequence starting at column  $l_i$  and ending at column  $r_i$ . Moreover two sequences  $C_i$  and  $C_{i+1}$  share at least 4 columns by the previous paragraph. Let us assume that in addition  $r_{i+1} > r_i$  (otherwise discard  $C_i$  with  $r_i = r_{i+1}$  and rename  $C_i$ 's and repeat until all  $r_i$ 's are distinct). We observe that after removal and renaming of  $C_i$ 's the condition  $l_{i+1} \leq r_i - 3$  still holds. So let us take a decreasing sequence contained in  $C_1$  starting at column  $l_1$  and ending at column  $r_1$ . Add to this all decreasing

sequences from  $C_{i+1}$ ,  $i \geq 1$  starting at  $r_i - 3$  ending at  $r_i$ . This gives us a desired good set of size at least  $2k + 1$ .

□

This finishes the proof that  $S_k$  is a canonical antichain for bichain graphs. An analogous proof shows that an antichain  $S_k^*$  obtained from  $S_k$  by replacing one part of bipartition by a clique is an antichain for split permutation graphs. This result follows from the following remarks:

- $S_k^*$  is an antichain (in the argument for  $S_k$  above replace  $2K_2$  by  $P_4$ )
- $Z_{n,k}^*$  is a  $k$ -letter graph (in Lemma 7.3.1 add pairs of letters corresponding to even columns to the set of connection rules)
- Any good set of size  $2k$  now corresponds to  $S_k^*$ , hence the same argument as in Lemma 7.3.3 shows that canonically prime split graphs (defined analogously to the bipartite case) embed into  $Z_{n,2k+1}^*$

## 7.4 Conclusion

In this section we identified a minimal class of unbounded clique-width and showed that it contains a canonical antichain with respect to labelled induced subgraph relation. Extending this result to the construction of a canonical antichain with respect to induced subgraph relation remains an open question.

It is an interesting and challenging research direction to further explore minimal classes of unbounded clique-width and obtain some general structural results of such classes with ultimate goal of identifying all such classes. Proving that all such classes are not well-quasi-ordered would then settle the conjecture of Daligault et al. [2010].

# Bibliography

- M. H. Albert and M. D. Atkinson. Simple permutations and pattern restricted permutations. *Discrete Mathematics*, 300(1):1–15, 2005.
- M. H. Albert and V. Vatter. Generating and enumerating 321-avoiding and skew-merged simple permutations. *The Electronic Journal of Combinatorics*, 20: Research paper 44, 11 pp, 2013.
- M. H. Albert, N. Ruškuc, and V. Vatter. Inflations of geometric grid classes of permutations. *Israel Journal of Mathematics*, to appear, 2014.
- V. E. Alekseev. Range of values of entropy of hereditary classes of graphs. *Diskretnaya Matematika*, 4(2):148–157, 1992. In Russian; translation in *Discrete Mathematics and Applications* 3(2):191–199, 1993.
- P. Allen, V. Lozin, and M. Rao. Clique-width and the speed of hereditary properties. *Electronic Journal of Combinatorics*, 16(1):Research Paper 35, 11 pp, 2009.
- N. Alon, J. Balogh, B. Bollobás, and R. Morris. The structure of almost all graphs in a hereditary property. *Journal of Combinatorial Theory, Series B*, 101(2):85–110, 2011.
- S. Alstrup and T. Rauhe. Small induced-universal graphs and compact implicit graph representations. *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 53–62, 2002.
- C. Arbib and R. Mosca. Polynomial algorithms for special cases of the balanced complete bipartite subgraph problem. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 39:3–22, 1999.
- S. Arnborg and A. Proskurowski. Linear time algorithms for NP-hard problems restricted to partial  $k$ -trees. *Discrete Applied Mathematics*, 23(1):11–24, 1989.

- M. D. Atkinson, M. M. Murphy, and N. Ruškuc. Partially well-ordered closed sets of permutations. *Order*, 19(2):101–113, 2002.
- A. Atminas and V. V. Lozin. Labelled induced subgraphs and well-quasi-ordering. *Order*, to appear, 2014.
- A. Atminas, R. Brignall, N. Korpelainen, V. Lozin, and V. Vatter. Well-quasi-order for permutation graphs omitting a path and a clique. *arXiv preprint arXiv:1312.5907*, 2013a.
- A. Atminas, V. Lozin, and M. Moshkov. Deciding wqo for factorial languages. *Lecture Notes in Computer Science*, 7810:68–79, 2013b.
- A. Atminas, V. V. Lozin, S. Kitaev, and A. Valyuzhenich. Universal graphs and universal permutations. *Discrete Mathematics, Algorithms and Applications*, 5(4), 2013c.
- A. Atminas, A. Collins, J. Foniok, and V. V. Lozin. Deciding the Bell number for hereditary graph properties. *arXiv preprint arXiv:1405.3118*, 2014a.
- A. Atminas, V. V. Lozin, and I. Razgon. Well-quasi-ordering, tree-width and sub-quadratic properties of graphs. *arXiv preprint arXiv:1410.3260*, 2014b.
- G. Bacsó and Z. Tuza. A characterization of graphs without long induced paths. *Journal of graph theory*, 14(4):455–464, 1990.
- J. Balogh, B. Bollobás, and D. Weinreich. The speed of hereditary properties of graphs. *Journal of Combinatorial Theory, Series B*, 79(2):131–156, 2000.
- J. Balogh, B. Bollobás, and D. Weinreich. The penultimate rate of growth for graph properties. *European Journal of Combinatorics*, 22(3):277–289, 2001.
- J. Balogh, B. Bollobás, and D. Weinreich. A jump to the Bell number for hereditary graph properties. *Journal of Combinatorial Theory, Series B*, 95(1):29–48, 2005.
- C. Benzaken, P. L. Hammer, and D. de Werra. Split graphs of Dilworth number 2. *Discrete mathematics*, 55(2):123–127, 1985.
- D. Binkle-Raible, H. Fernau, S. Gaspers, and M. Liedloff. Exact exponential-time algorithms for finding bicliques. *Information Processing Letters*, 111(2):64–67, 2010.
- H. L. Bodlaender. On linear time minor tests with depth-first search. *Journal of Algorithms*, 14(1):1–23, 1993.



- A. Brandstädt, T. Klemmt, and S. Mahfud.  $P_6$ - and triangle-free graphs revisited: structure and bounded clique-width. *Discrete Mathematics and Theoretical Computer Science*, 8(1):173–188, 2006.
- R. Brignall. Grid classes and partial well order. *Journal of Combinatorial Theory, Series A*, 119(1):99–116, 2012.
- R. Brignall, N. Ruškuc, and V. Vatter. Simple permutations: decidability and unavoidable substructures. *Theoretical Computer Science*, 391(1):150–163, 2008.
- H. Broersma, P. A. Golovach, D. Paulusma, and J. Song. Updating the complexity status of coloring graphs without a fixed induced linear forest. *Theoretical Computer Science*, 414(1):9–19, 2012.
- A. A. Bulatov and D. Marx. Constraint satisfaction parameterized by solution size. *SIAM Journal on Computing*, 43(2):573–616, 2014.
- S. Butler. Induced-universal graphs for graphs with bounded maximum degree. *Graphs and Combinatorics*, 25(4):461–468, 2009.
- Y. Chen, M. Thurley, and M. Weyer. Understanding the complexity of induced subgraph isomorphisms. *Proceedings of the 35th international colloquium on Automata, Languages and Programming, Part I*, pages 587–596, 2008.
- G. L. Cherlin and B. J. Latka. Minimal antichains in well-founded quasi-orders with an application to tournaments. *Journal of Combinatorial Theory, Series B*, 80(2):258–276, 2000.
- F. R. Chung. Universal graphs and induced-universal graphs. *Journal of Graph Theory*, 14(4):443–454, 1990.
- V. Chvátal and P. L. Hammer. Aggregation of inequalities in integer programming. *Annals of Discrete Mathematics*, 1:145–162, 1977.
- B. Courcelle, J. Engelfriet, and G. Rozenberg. Handle-rewriting hypergraph grammars. *Journal of Computer and System Sciences*, 46(2):218–270, 1993.
- B. Courcelle, J. A. Makowsky, and U. Rotics. Linear time solvable optimization problems on graphs of bounded clique-width. *Theory of Computing Systems*, 33(2):125–150, 2000.
- M. Crochemore, F. Mignosi, and A. Restivo. Automata and forbidden words. *Information Processing Letters*, 67(3):111–117, 1998.

- J. Daligault, M. Rao, and S. Thomassé. Well-quasi-order of relabel functions. *Order*, 27(3):301–315, 2010.
- P. Damaschke. Induced subgraphs and well-quasi-ordering. *Journal of graph theory*, 14(4):427–435, 1990.
- A. de Luca and S. Varricchio. Well quasi-orders and regular languages. *Acta Informatica*, 31(6):539–557, 1994.
- G. Ding. Subgraphs and well-quasi-ordering. *Journal of Graph Theory*, 16(5):489–502, 1992.
- G. Ding. Stable sets versus independent sets. *Discrete Mathematics*, 117(1):73–87, 1993.
- G. Ding. On canonical antichains. *Discrete Mathematics*, 309(5):1123–1134, 2009.
- J. Dong. Some results on graphs without long induced paths. *Journal of Graph Theory*, 22(1):23–28, 1996.
- P. Erdős, D. J. Kleitman, and B. L. Rothschild. Asymptotic enumeration of  $K_n$ -free graphs. *Citeseer*, 1976.
- P. Erdős, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs and Combinatorics*, 2(1):113–121, 1986.
- L. Esperet, A. Labourel, and P. Ochem. On induced-universal graphs for the class of bounded-degree graphs. *Information Processing Letters*, 108(5):255–260, 2008.
- M. R. Fellows and M. A. Langston. On search decision and the efficiency of polynomial-time algorithms. *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 501–512, 1989.
- A. Finkel and P. Schnoebelen. Well-structured transition systems everywhere! *Theoretical Computer Science*, 256(1):63–92, 2001.
- S. Földes and P. L. Hammer. Split graphs. *Proceedings of the Eighth Southeastern Conference on Combinatorics, Graph Theory and Computing*, 19:311–315, 1977.
- J.-L. Fouquet, V. Giakoumakis, and J.-M. Vanherpe. Bipartite graphs totally decomposable by canonical decomposition. *International Journal of Foundations of Computer Science*, 10(04):513–533, 1999.

- Z. Füredi and P. Komjáth. Nonexistence of universal graphs without some trees. *Combinatorica*, 17(2):163–171, 1997.
- T. Gallai. Transitiv orientierbare Graphen. *Acta Mathematica Hungarica*, 18(1):25–66, 1967.
- J. Gustedt. Finiteness theorems for graphs and posets obtained by compositions. *Order*, 15(3):203–220, 1998.
- A. Gyárfás and J. Lehel. A Helly-type problem in trees. *Combinatorial theory and its applications*, 4:571–584, 1970.
- P. L. Hammer and A. K. Kelmans. On universal threshold graphs. *Combinatorics, probability and computing*, 3(03):327–344, 1994.
- G. Higman. Ordering by divisibility of abstract algebras. *Proceedings London Mathematical Society*, 2:326–336, 1952.
- N. Hine and J. Oxley. When excluding one matroid prevents infinite antichains. *Advances in Applied Mathematics*, 45(1):74–76, 2010.
- S. Huczynska and V. Vatter. Grid classes and the Fibonacci dichotomy for restricted permutations. *Electronic Journal of Combinatorics*, 13:Research paper 54, 14 pp., 2006.
- D. S. Johnson. The NP-completeness column: An ongoing guide. *Journal of Algorithms*, 8(3):438–448, 1987.
- M. Kamiński, V. Lozin, and M. Milanič. Recent developments on graphs of bounded clique-width. *Discrete applied mathematics*, 157:2747–2761, 2009.
- S. Kannan, M. Naor, and S. Rudich. Implicit representation of graphs. *SIAM Journal on Discrete Mathematics*, 5(4):596–603, 1992.
- T. Kővári, V. T. Sós, and P. Turán. On a problem of K. Zarankiewicz. *Colloquium Mathematicum*, 3:50–57, 1954.
- S. Kitaev. Patterns in permutations and words. *Springer-Verlag*, 2011.
- P. G. Kolaitis, H.-J. Prömel, and B. L. Rothschild.  $K_{l+1}$ -free graphs: asymptotic structure and a 0–1 law. *Transactions of the American Mathematical Society*, 303(2):637–671, 1987.

- P. Komjáth and J. Pach. Universal graphs without large bipartite subgraphs. *Mathematika*, 31(02):282–290, 1984.
- P. Komjáth, A. H. Mekler, and J. Pach. Some universal graphs. *Israel Journal of Mathematics*, 64(2):158–168, 1988.
- D. Korobitsyn. On the complexity of determining the domination number in monogenic classes of graphs. *Diskretnaya Matematika*, 2(3):90–96 (in Russian), 1990.
- N. Korpelainen. Boundary properties of graphs. *PhD thesis, University of Warwick*, 2012.
- N. Korpelainen and V. Lozin. Two forbidden induced subgraphs and well-quasi-ordering. *Discrete Mathematics*, 311(16):1813–1822, 2011a.
- N. Korpelainen and V. V. Lozin. Bipartite induced subgraphs and well-quasi-ordering. *Journal of Graph Theory*, 67(3):235–249, 2011b.
- N. Korpelainen, V. V. Lozin, and C. Mayhill. Split permutation graphs. *Graphs and Combinatorics*, 30(3):633–646, 2014.
- D. Král, J. Kratochvíl, Z. Tuza, and G. J. Woeginger. Complexity of coloring graphs without forbidden induced subgraphs. *Proceedings of the 27th International Workshop on Graph-Theoretic Concepts in Computer Science*.
- J. B. Kruskal. Well-quasi-ordering, the tree theorem, and Vazsonyi’s conjecture. *Transactions of the American Mathematical Society*, 95:210–225, 1960.
- J. B. Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A*, 13(3):297–305, 1972.
- D. Kühn and D. Osthus. Induced subdivisions in  $K_{s,s}$ -free graphs of large average degree. *Combinatorica*, 24(3):287–304, 2004.
- V. B. Le, B. Randerath, and I. Schiermeyer. On the complexity of 4-coloring graphs without long induced paths. *Theoretical Computer Science*, 389(1):330–335, 2007.
- V. Lozin and D. Rautenbach. Some results on graphs without long induced paths. *Information Processing Letters*, 88(4):167–171, 2003.
- V. V. Lozin. On minimal universal graphs for hereditary classes. *Diskretnaya Matematika*, 9(2):106–115, 1997. In Russian, translation in *Discrete Mathematics and Applications* 7(3): 295–304, 1997.

- V. V. Lozin. Minimal classes of graphs of unbounded clique-width. *Annals of Combinatorics*, 15(4):707–722, 2011.
- V. V. Lozin and C. Mayhill. Canonical antichains of unit interval and bipartite permutation graphs. *Order*, 28(3):513–522, 2011.
- V. V. Lozin and G. Rudolf. Minimal universal bipartite graphs. *Ars Combinatoria*, 84:345–356, 2007.
- J. Moon. On minimal  $n$ -universal graphs. 7(01):32–33, 1965.
- M. M. Murphy. Restricted permutations, antichains, atomic classes, and stack sorting. *PhD thesis, University of St. Andrews*, 2002.
- M. M. Murphy and V. Vatter. Profile classes and partial well-order for permutations. *Electronic Journal of Combinatorics*, 9(2), 2003.
- S. Norine, P. Seymour, R. Thomas, and P. Wollan. Proper minor-closed families are small. *Journal of Combinatorial Theory, Series B*, 96(5):754–757, 2006.
- M. Petkovšek. Letter graphs and well-quasi-order by induced subgraphs. *Discrete mathematics*, 244(1):375–388, 2002.
- M. Pouzet. Un bel ordre d’abritement et ses rapports avec les bornes d’une multirelation. *Comptes Rendus de l’Académie des Sciences Paris Série A-B*, 274: A1677–A1680, 1972.
- H. J. Prömel and A. Steger. Excluding induced subgraphs: quadrilaterals. *Random Structures & Algorithms*, 2(1):55–71, 1991.
- H. J. Prömel and A. Steger. Excluding induced subgraphs III: a general asymptotic. *Random Structures & Algorithms*, 3(1):19–31, 1992.
- H. J. Prömel and A. Steger. Excluding induced subgraphs II: extremal graphs. *Discrete Applied Mathematics*, 44(1):283–294, 1993.
- R. Rado. Universal graphs and universal functions. *Acta Arithmetica*, 9:331–340, 1964.
- B. Randerath and I. Schiermeyer. 3-colorability in  $P$  for  $P_6$ -free graphs. *Discrete Applied Mathematics*, 2(3):299–313, 2004.
- N. Robertson and P. D. Seymour. Graph minors. XX. Wagner’s conjecture. *Journal of Combinatorial Theory, Series B*, 92(2):325–357, 2004.

- N. Robertson, P. Seymour, and R. Thomas. Quickly excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 62(2):323–348, 1994.
- E. R. Scheinerman and J. Zito. On the size of hereditary classes of graphs. *Journal of Combinatorial Theory, Series B*, 61(1):16–39, 1994.
- D. A. Spielman and M. Bóna. An infinite antichain of permutations. *The electronic journal of combinatorics*, 7(2):1–4, 2000.
- V. Vatter. Small permutation classes. *Proceedings of the London Mathematical Society*, 103(5):879–921, 2011.
- V. Vatter and S. Waton. On partial well-order for monotone grid classes of permutations. *Order*, 28(2):193–199, 2011.
- G. J. Woeginger, J. Sgall, et al. The complexity of coloring graphs without long induced paths. *Acta Cybern.*, 15(1):107–117, 2001.
- M. Yannakakis. The complexity of the partial order dimension problem. *SIAM Journal on Algebraic Discrete Methods*, 3(3):351–358, 1982.
- I. E. Zverovich. A solution to a problem of Jacobson, Kézdy and Lehel. *Graphs and Combinatorics*, 20(4):571–577, 2004.